UNIVERSITÀ DEGLI STUDI DI PISA

MASTER'S DEGREE IN MATHEMATICS

Volatility and Dispersion strategies in Finance

MASTER'S THESIS

Candidate
Federico Borghese

Supervisor
Maurizio Pratelli
Università di Pisa

Supervisor
Antoine Garaïalde
J.P. Morgan

ANNO ACCADEMICO 2015/2016
October 14, 2016

Università degli Studi di Pisa

Master's Degree in Mathematics

Volatility and Dispersion strategies in Finance

Candidate
Federico Borghese[1]

Federico Borghese

Supervisor
Maurizio Pratelli
Università di Pisa

Supervisor
Antoine Garaïalde
J.P. Morgan

Antoine Garaïalde

October 14, 2016

[1]Equity Derivatives Structuring at J.P. Morgan - Università di Pisa - Scuola Normale Superiore - M2 Probabilités et Finance (Paris 6 & École Polytechnique).
Contact: fed.borghese@gmail.com

Contents

Chapter 1

J.P. Morgan Equity Derivatives Structuring Group

1.1 J.P. Morgan

J.P. Morgan is one of the most respected financial services firms in the world, serving governments, corporations and institutions in over 100 countries. The firm is a leader in investment banking, financial services for consumers and small businesses, commercial banking, financial transaction processing and asset management.
With a history dating back over 200 years, J.P. Morgan is one of the oldest financial institutions in the United States. It is the largest bank in the US, and the world's sixth largest bank by total assets, with total assets of US$2.4 trillion. It employs more than $235,000$ people worldwide.
The Firm and its Foundation give approximately $200 million annually to non-profit organizations around the world. J.P. Morgan also leads volunteer service activities for employees in local communities.

1.2 Equity Derivatives Structuring

Equity Derivatives Structuring focuses on developing alternative payoff profiles for a wide range of investors, running from highly sophisticated institutional investors (Hedge Funds, Asset Managers, Pension Funds, Insurance Companies) to less sophisticated ones (Retail). The Team designs and prices both new and existing structured products in the core equity derivatives space. There is no precise definition of what alternative payoffs are; everything which is not a easily accessed by vanilla options and futures on common asset classes can be regarded as an alternative payoff.

The product offering is very wide. It includes new derivative ideas inspired from research, existing derivatives and fully tailored solutions for the clients which are designed according to their return expectations, their risk appetite and their views on the market. For example, the Team offers sophisticated institutional investors ideas to trade volatility, correlation and skew.
The business has considerably changed after the 2008 financial crisis. The regulator is not in favour of complexity and the range of the products has significantly shrunk. However, the range of possible underlyings has considerably widened and the interest from clients stays strong. J.P. Morgan develops proprietary indices which implement systematic strategies on which simple derivative products can be written, achieving simplicity and innovation at the same time.

1.2.1 A bridge between Trading and Sales

Equity Derivatives Structurers have the important task to price their products. The Black-Scholes model and its variations, which are widely used for pricing structured products, can't keep in consideration all the possible factors influencing the market evolution. For example,

- Volatility is not deterministic, but stochastic. The market is not *complete*, i.e. it is not possible to perfectly replicate derivatives using the underlying only.
- The are jumps in the spot price evolution. Often these jumps are due to news such as macroeconomics news, central bank policy, political news, release of a new product, quarterly reports etc. Jumps are highly unpredictable and always set big challenges to hedging.
- There are transaction costs to trading the underlying and there is a bid-offer spread.
- Not every underlying and vanilla option is available in the market.

Structurers' skills include determining these *extra-model* risks and incorporate them in the price.
The products are sold by Sales people, who always need the lowest possible price to beat the competition and secure the trade with the client. After the sale, the Trading Desk has to hedge the product in order to neutralize the bank's position and avoid any losses. Therefore the Trading Desk wants to incorporate in the price all the possible hedging risks. The goal of the price computed by the Structuring Desk is to make a good compromise between Sales and Trading.

Moreover, in the product development process, the Structuring Desk needs to know both the Sales and the Trading points of view. They need to know the clients, the type of products which are popular in the market, they have to find attractive payoffs and underlyings that can be advertised. But they also have to think about the hedging strategies, the potential sources of risk, the sensitivities, the status of the traders' books.

Chapter 2

Preparatory concepts

This brief chapter recalls a few mathematical tools that will be used throughout this document. We don't report the basics of mathematical finance, which can be found in [2] and [3].

2.1 European Payoff replication with vanilla options

This result is sometimes known as Carr's formula. Let $G : \mathbb{R}^+ \to \mathbb{R}$ a twice differentiable function (can be relaxed to G being the difference of two convex functions or even more than that). Consider a contract on an underlying S which pays at maturity T the amount $G(S_T)$. This is what is called a *European* payoff: the final value of the contract depends only on the final price of the underlying, and the contract cannot be unwound before maturity.
Suppose we have a liquid market of Call and Put Options on the underlying with the same maturity T, with quoted prices $Call(K)$ for every strike $K \geq S^*$ and $Put(K)$ for every strike $K \leq S^*$.
Then we can perfectly replicate the European Payoff $G(S_T)$ using only a zero-coupon bond, a forward contract on the underlying and the Calls and Puts in the following manner:

$$\boxed{\begin{aligned} G(S_T) = G(S^*) + G'(S^*)(S_T - S^*) + \\ + \int_0^{S^*} G''(K)(K - S_T)_+ dK + \int_{S^*}^{\infty} G''(K)(S_T - K)_+ dK \end{aligned}} \tag{2.1}$$

As a consequence, the only allowable price for the contract in order to prevent any arbitrage is

$$\begin{aligned}\text{Contract price} = G(S^*)B(0,T) + G'(S^*)(\text{price of a forward struck at } S^*)+ \\ + \int_0^{S^*} G"(K)Put(K)dK + \int_{S^*}^{\infty} G"(K)Call(K)dK\end{aligned} \tag{2.2}$$

where $B(0,T)$ is the price of a zero-coupon bond with maturity T.

2.1.1 Proof

Proof. Assume that G is twice differentiable[1]. By the integration by parts formula, we have

$$\begin{aligned}\int_{S^*}^{+\infty} G"(K)(x-K)_+ dK &= [G'(K)(x-K)_+]_{S^*}^{+\infty} + \int_{S^*}^{+\infty} G'(K)1_{\{x\geq K\}}dK \\ \int_0^{S^*} G"(K)(K-x)_+ dK &= [G'(K)(K-x)_+]_0^{S^*} - \int_0^{S^*} G'(K)1_{\{x\leq K\}}dK\end{aligned} \tag{2.3}$$

The terms in brackets evaluated at $K=0$ and $K=+\infty$ are equal to zero, so we can rewrite

$$\begin{aligned}\int_{S^*}^{+\infty} G"(K)(x-K)_+ dK &= -G'(S^*)(x-S^*)_+ + (G(x\vee S^*) - G(S^*)) \\ \int_0^{S^*} G"(K)(K-x)_+ dK &= G'(S^*)(S^*-x)_+ - (G(S^*) - G(x\wedge S^*))\end{aligned} \tag{2.4}$$

By summing the previous two equations, we get:

$$\begin{aligned}\int_0^{S^*} G"(K)(K-x)_+ dK + \int_{S^*}^{+\infty} G"(K)(x-K)_+ dK = \\ = G'(S^*)(S^*-x) - G(S^*) + G(x)\end{aligned} \tag{2.5}$$

Reordering the terms and evaluating at $x = S_T$ we find back equation (2.1). □

[1]The proof is the same in the case of G being the difference of convex functions. In this case G has a left derivative and its second derivative in the sense of distributions (a Radon measure) satisfies the integration by parts by definition.

2.2 Ito-Tanaka's formula

Ito's formula, one of the most useful tools in mathematical finance, is only valid for C^2 functions. In the market, however, we very often deal with functions which are not twice differentiable. Even the simplest derivative product, a Call Option, has a payoff which is not C^2.
This regularity hypothesis is very often neglected by practitioners, because in many cases a "cheating" Ito's formula provides the good result. In this section we provide rigorous theorems which allow us to deal with functions belonging to a wider regularity class. Ito-Tanaka's formula can be applied to all functions which are difference of convex functions.

Let $f : \mathbb{R} \to \mathbb{R}$ a convex function, and let f'_- be its left derivative. Being f convex, the second derivative f" in the sense of distributions is a positive Radon measure. Let (X_t) be a continuous semimartingale with quadratic variation $[X, X]_t$ and take $a \in \mathbb{R}$; we define a stochastic process $(L_t^a)_{t\geq 0}$, the *Local Time of* X *at* a:

$$L_t^a = \lim_{\varepsilon \searrow 0} \frac{1}{\varepsilon} \int_0^t 1_{[a,a+\varepsilon)}(X_s) d[X, X]_s$$

The Local Time is an increasing, non-negative process, null at 0. L_t^a can be thought as the time spent by the semimartingale at point a in the time period $[0, t]$, measured in the natural time scale of the semimartingale. For further details see [1].

Theorem 2.1. *(Ito-Tanaka's formula) Take* $f : \mathbb{R} \to \mathbb{R}$ *convex and* $(X_t)_{t\in[0,T]}$ *a continuous semimartingale. Then*

$$f(X_t) - f(X_0) = \int_0^t f'_-(X_s)dX_s + \frac{1}{2}\int_{\mathbb{R}} L_t^a f"(da)$$

where L_t^a *is the Local Time of the semimartingale and* $f"(da)$ *is the Radon measure associated with the second derivative of* f.

In many cases in finance, the measure f" is absolutely continuous with respect to Lebesgue measure. In this simpler case, we can use the Occupation Times Formula to handle the Local Time term in the previous formula:

Theorem 2.2. *(Occupation Times) Take* (X_t) *a semimartingale. Almost surely,* $\forall T > 0$ *and for all* $F : \mathbb{R} \to \mathbb{R}$ *positive Borelian function,*

$$\int_{-\infty}^{+\infty} L_t^a F(a)da = \int_0^T F(X_s)d[X, X]_s$$

where $[X, X]_s$ *is the quadratic variation of the semimartingale.*

Being the two previous formulas linear in f, obviously they can be extended to all the functions that can be written as a difference of convex function. Combining the two formulas we can state the following extended Ito's Formula:

Theorem 2.3. *Let $(X_t)_{t\in[0,T]}$ be a continuous semimartingale. Let $f : \mathbb{R} \to \mathbb{R}$ be the difference of two convex functions. Assume that the second derivative of f (which is a signed Radon measure) is absolutely continuous with respect to Lebesgue measure, with its Radon-Nikodym derivative denoted $f"(x)$. Then*

$$f(X_t) - f(X_0) = \int_0^t f'_-(X_s)dX_s + \frac{1}{2}\int_0^t f"(X_s)d[X,X]_s \tag{2.6}$$

2.3 Gamma P&L

When an investment bank sells a derivative product, its trading desk usually Delta-Hedges it. Therefore, what really matters for the P&L of the bank is not the payoff of the product itself, but *the P&L of the Delta-Hedged product.* Let $g(S_T)$ be an European payoff on the underlying S_t, with $v_t = v(t, S_t)$ being the fair price of the product at time t. The Theta, Delta and Gamma of the product are

$$\theta_t = \frac{\partial v(t,x)}{\partial t}(x = S_t)\,; \qquad \delta_t = \frac{\partial v(t,x)}{\partial x}(x = S_t)\,; \qquad \Gamma_t = \frac{\partial^2 v(t,x)}{\partial x^2}(x = S_t) \tag{2.7}$$

Assume for simplicity the existence of a risk-free rate r. The Delta-Hedge portfolio V_t is a self-financed portfolio with $V_0 = v_0$ equal to the option premium and

$$dV_t = \delta_t dS_t + (V_t - \delta_t S_t)rdt \tag{2.8}$$

Let us work in the Black-Scholes framework, where

$$\frac{dS_t}{S_t} = rdt + \sigma dW_t \tag{2.9}$$

where W is a Brownian Motion under the risk-neutral measure. The position of the trader having sold the product is then $V_t - v_t$.
By Black-Scholes equation, we have

$$\theta_t + rS_t\delta_t + \frac{1}{2}\Gamma_t\sigma^2 S_t^2 - rv = 0 \tag{2.10}$$

Let Δt be a short time period, say one day. Over this period, by Taylor expansion on v, the changes in the option value is

$$\Delta v \approx \theta \Delta t + \delta \Delta S + \frac{1}{2}\Gamma(\Delta S)^2 \tag{2.11}$$

By equation (2.10), we have

$$\begin{aligned} \Delta V &= \delta \Delta S + (V - \delta S) r \Delta t = \\ &= \delta \Delta S + (v - \delta S) r \Delta t + (V - v) r \Delta t = \\ &= \delta \Delta S + \left(\theta + \frac{1}{2}\Gamma \sigma^2 S^2\right) \Delta t + (V - v) r \Delta t \end{aligned} \tag{2.12}$$

Combining the previous equations we find the total daily Gamma P&L of the trader:

$$P\&L^{\Gamma} = \Delta(V_t - v_t) \approx \frac{1}{2}\Gamma_t S_t^2 \left[\sigma^2 \Delta t - \left(\frac{\Delta S_t}{S_t}\right)^2\right] + (V - v) r \Delta t \tag{2.13}$$

The previous equation can be interpreted in terms of the discounted P&L:

$$\text{discounted } P\&L^{\Gamma} = \Delta\left(e^{-rt}(V_t - v_t)\right) \approx e^{-rt}\frac{1}{2}\Gamma_t S_t^2 \left[\sigma^2 \Delta t - \left(\frac{\Delta S_t}{S_t}\right)^2\right] \tag{2.14}$$

Assuming that the replication is not too far from the fair value of the option, i.e. $v_t \approx V_t$ we can write the simplest form of the Gamma P&L:

$$\boxed{P\&L^{\Gamma} = \Delta(V_t - v_t) \approx \frac{1}{2}\Gamma_t S_t^2 \left[\sigma^2 \Delta t - \left(\frac{\Delta S_t}{S_t}\right)^2\right]} \tag{2.15}$$

2.4 Basket of n stocks

There are many ways to define a *Basket* of stocks. All the variations have in common that a Basket is a strategy which invests its wealth among many different stocks, but the ways to choose the weights and the rebalancing can lead to very different results.
Let $S_t^1, ..., S_t^n$ be the prices of n stocks at time t. The daily returns of the stocks are $\dfrac{\Delta S_t^i}{S_t^i} = \dfrac{S_{t+1}^i}{S_t^i} - 1$. Let $w_t^1, ..., w_t^n$ be the weights of the stocks, with $\sum_i w_t^i = 1$.

Definition 2.4. A **Basket** of the n stocks, with weights w_t^i, is a portfolio B which invests every day $w_t^i\%$ of the wealth in the i-th stock. That is,

$$\frac{\Delta B_t}{B_t} = \sum_{i=1}^{n} w_t^i \frac{\Delta S_t^i}{S_t^i} \tag{2.16}$$

A large number of commonly traded baskets fall into this framework:

- **Dynamic Basket:** A strategy with constant weights. It is called *dynamic* because the number of shares NOSH_t^i held in the basket changes every day and a daily dynamic rebalancing is needed. If $w^1, ..., w^n$ are the fixed weights of the stocks, the Basket return is

$$\frac{\Delta B_t}{B_t} = \sum_{i=1}^{n} w^i \frac{\Delta S_t^i}{S_t^i} \tag{2.17}$$

$$\Rightarrow \quad \Delta B_t = \sum_{i=1}^{n} \frac{w^i B_t}{S_t^i} \Delta S_t^i \tag{2.18}$$

Therefore $\text{NOSH}_t^i = \dfrac{w^i B_t}{S_t^i}$.

- **Static Basket:** A strategy V which has a fixed number of shares NOSH_i in each stock. It is called *static* because there is no rebalancing. Usually the Basket value is fixed to 1 at time zero, and initial weights $w_0^1, ..., w_0^n$ are provided rather than the NOSH. The initial weights and the NOSH are linked by the following relation:

$$w_0^i = \frac{\text{NOSH}_i \cdot S_0^i}{V_0} = \text{NOSH}_i \cdot S_0^i \tag{2.19}$$

The value of the basket is:

$$V_t = \sum_{i=1}^{n} \text{NOSH}_i S_t^i = \sum_{i=1}^{n} w_0^i \frac{S_t^i}{S_0^i} \tag{2.20}$$

In this case $\frac{\Delta V_t}{V_t} = \frac{1}{V_t} \sum_{i=1}^{n} \text{NOSH}^i \Delta S_t^i$, therefore

$$\frac{\Delta V_t}{V_t} = \sum_{i=1}^{n} \frac{\text{NOSH}_i \cdot S_t^i}{V_t} \frac{\Delta S_t^i}{S_t^i} \tag{2.21}$$

The weights (which have unit sum) are:

$$w_t^i = \frac{\text{NOSH}_i \cdot S_t^i}{V_t} = \frac{w_0^i \cdot S_t^i}{V_t \cdot S_0^i}$$

The weights change every day and the best performing stocks see their weight increase. Observe that, after a long period of time, stock prices can go very far from the initial price, so the strategy will become more and more similar to holding the best performing stocks (and less dependent on correlation).
Most of the times, when people say "*a basket 60% SX5E, 30% UKX and 10% SMI*" they don't mean that the weights are fixed at 60%,30% and 10% but rather that the NOSH is fixed. The value of the basket they refer to is

$$V_t = 60\% \frac{\text{SX5E}_t}{\text{SX5E}_0} + 30\% \frac{\text{UKX}_t}{\text{UKX}_0} + 10\% \frac{\text{SMI}_t}{\text{SMI}_0} \tag{2.22}$$

- **Arithmetic baskets**: they are a special case of the above. An arithmetic basket starting at time 0 is defined as

$$B_t = \frac{1}{n} \sum_{i=1}^{n} \frac{S_t^i}{S_0^i} \tag{2.23}$$

Therefore, $\dfrac{\Delta B_t}{B_t} = \displaystyle\sum_{i=1}^{n} \frac{S_t^i}{n S_0^i B_t} \frac{\Delta S_t^i}{S_t^i}$, so the weights are $w_t^i = \dfrac{S_t^i}{n S_0^i B_t}$

- The Euro Stoxx 50 (SX5E Index). This is a basket of the 50 largest European companies. It rebalances every 3 months, and between two rebalancing dates it holds a fixed number of shares. Let k be a rebalancing date and w_k^i the weight of stock i for $i = 1, ..., 50$. Then

$$\frac{\Delta \text{SX5E}_k}{\text{SX5E}_k} = \sum_{i=1}^{50} w_k^i \frac{\Delta S_k^i}{S_k^i} \tag{2.24}$$

The number of shares of stock i is then $\text{NOSH}^i = \dfrac{w_k^i \text{SX5E}_k}{S_k^i}$. Hence, the effective weights of the Euro Stoxx 50 are

$$w_t^i = \frac{w_k^i \cdot \text{SX5E}_k \cdot S_t^i}{S_k^i \cdot \text{SX5E}_t} \tag{2.25}$$

Observe that these weights, differently from the previous cases, remain meaningful as time passes because of the rebalancing. Over three months, $\dfrac{\mathrm{SX5E}_k \cdot S_t^i}{S_k^i \cdot \mathrm{SX5E}_t}$ rarely becomes significantly different than 1, so, with a good approximation, we could say that the EuroStoxx has constant weights $w_t^i = w_k^i$ for every 3-month period.

Chapter 3

Volatility products

3.1 Introduction

Volatility is one of the most important quantities in financial markets, given its key role in determining the price and the hedge of derivatives, in evaluating an asset's performance (e.g. Sharpe Ratio), in designing investment strategies and Smart Beta Indices.

Volatility is a measure of the way a stochastic process moves around its trend. A very smooth stochastic process, for example the value of a strategy which invests in money market instruments has almost zero volatility. On the other hand, the price of a share of a company can be very volatile, especially in times of uncertainty and during market stress. The more volatile is a stock, the harder it is to predict its price in one month, the more expensive will be a Call or Put option on that stock. On the other hand, the most volatile assets offer the chances of the highest returns.

Volatility is a quantity which is not directly measurable. Within the classical Black-Scholes model, volatility is a parameter of the model and there are many good estimators of it. However there is evidence from data that volatility is not constant at all: it is not constant over time and even at a given time it varies depending on the strike and maturity of the option (*volatility smile*). Actually, *volatility does not exist*: in theory it should be the *instantaneous* rate at which a stochastic process is oscillating, which is impossible to measure or properly define, and volatility itself is a stochastic process.

While there are a variety of models which introduce a stochastic diffusion for the volatility (Dupire's local volatility, Heston's model and many other modern models), it is not in the scope of this document to treat them. The way we can treat volatility is the following: well defining *different types* of volatilities and understanding the meaning and the relationships between them.

3.1.1 Implied Volatility

Implied volatility is not a completely trivial quantity to understand. It can be seen as a forward-looking measure of how much the price *will* oscillate in the future. It is a quantity which is *implied by the market*, i.e. it is derived my the market price of products which depend on volatility. These products are Call and Put options, which in most cases are the most liquid derivative product available in the market. Therefore, implied volatility refers to:

- An underlying
- A particular strike
- A particular maturity

Definition 3.1. The implied volatility of the underlying S, at strike K and maturity T, is the volatility that, input to Black-Scholes pricing model, yields a theoretical price of a vanilla option with strike K and maturity T equal to the current market price of the option.

The way implied volatility can be seen is simply as a more convenient way to express a price. Currency prices are not the most expressive way to describe the price of an option: it is hard to compare a price of 7 USD for a 1Y at-the-money Call Option on the S&P 500 and a price of 150 EUR for a 2Y Quanto Straddle on the Nikkei 225. But it becomes meaningful to compare their implied volatilities which are expressed in the same units.

3.1.2 Realized Volatility

Realized volatility is a backwards-looking quantity, which measures how much the price has oscillated in a well specified time period.

Definition 3.2. The daily **Realized Variance** of an Underlying S in the period $[t, T]$ (expressed in number of days) is the following quantity:

$$\sigma^2_{t,T} := \frac{252}{T-t} \sum_{u=t+1}^{T} \ln\left(\frac{S_u}{S_{u-1}}\right)^2 \tag{3.1}$$

The daily **Realized Volatility** is the square root of the Realized Variance, i.e.

$$\sigma_{t,T} := \sqrt{\frac{252}{T-t} \sum_{u=t+1}^{T} \ln\left(\frac{S_u}{S_{u-1}}\right)^2}$$

The length of the time period is expressed as a number of business days $(T-t)$ divided by the total number of business days in one year (by convention, 252).

Proposition 3.3. ***(Additivity property)****: The realised variance is additive on adjacent time periods. In formulas,*

$$(T_3 - T_1)\sigma^2_{T_1,T_3} = (T_2 - T_1)\sigma^2_{T_1,T_2} + (T_3 - T_2)\sigma^2_{T_2,T_3} \tag{3.2}$$

Proof. By definition (3.1), the right hand side is equal to

$$252 \sum_{u=T_1+1}^{T_2} \ln\left(\frac{S_u}{S_{u-1}}\right)^2 + 252 \sum_{u=T_2+1}^{T_3} \ln\left(\frac{S_u}{S_{u-1}}\right)^2$$

Combining the sums gives exactly the definition of $(T_3 - T_1)\sigma^2_{T_1,T_3}$. □

The are other possibilities for measuring the realised volatility. For example, instead of taking daily returns, we could introduce a *picking frequency* of, let's say, 5, and use 5-day returns in the previous formula. This is particularly appropriated when measuring the volatility of a basket of assets belonging to different geographical areas. We will come back to this topic in section 4.1.4.

Also, some people take different formulas for realised volatility, for example the standard deviation of the log-returns:

$$\sigma^2_{t,T} := \frac{252}{T-t-1} \sum_{u=t+1}^{T} \left[\ln\left(\frac{S_u}{S_{u-1}}\right) - \frac{1}{T-t} \sum_{v=t+1}^{T} \ln\left(\frac{S_v}{S_{v-1}}\right)\right]^2 \tag{3.3}$$

In most cases, the underlying price has a very small drift, so the two formulas give almost the same result. However, the latter formula excludes the drift and only measures the oscillating side of the price evolution. When computing the realised volatility in a hypothetical scenario where interest rates are 10%, the differences between the two formulas arise. A major inconvenience of this formula is that the additivity property is lost.

Realised volatility in a stochastic volatility model

If we assume a diffusive continuous model for the Underlying's price with stochastic volatility, that is

$$\frac{dS_t}{S_t} = \mu_t dt + \sigma_t dB_t \tag{3.4}$$

then it is natural to define the realized variance as

$$\sigma^2_{t,T} = \frac{1}{T-t}\int_t^T \sigma^2_u du$$

Actually, the definition given in (3.1) is a good approximation of the latter theoretical realized variance. In fact, if we assume the model (3.4), by Ito's formula we have

$$d\ln(S_t) = (\mu_t - \frac{1}{2}\sigma^2_t)dt + \sigma_t dW_t \tag{3.5}$$

Under suitable integrability assumptions, $\ln(S_t)$ is a continuous semimartingale. Therefore, by Theorems 1.8 and 1.18 in [1], we have a convergence in probability of

$$\sum_{i=1}^{N}(\ln(S_{i\frac{T}{N}}) - \ln(S_{(i-1)\frac{T}{N}}))^2 \xrightarrow{P} \langle \ln(S_t), \ln(S_t)\rangle = \int_0^T \sigma^2_t dt \tag{3.6}$$

for $N \to \infty$. However, this convergence holds for the time frame used to compute returns going to zero.

3.2 Variance Swaps

Definition 3.4. A **Variance Swap** is a forward contract on the annualized variance of the Underlying.

As any forward contract, the Variance Swap is an agreement between two counterparties to buy (or sell) at maturity an unknown quantity, in our case the future realized variance, at a certain strike price. Let K^2_{var} be the strike price of the Variance Swap and $\sigma^2_{0,T}$ the realized variance of the Underlying over the period $[0, T]$. Then the payoff at maturity of a Variance Swap expiring at time T is

$$N(\sigma^2_{0,T} - K^2_{var})$$

where N is the notional of the contract.
In most cases, the Variance Swap notional is expressed in the form of a *Vega Notional*. The payoff is

$$(\text{Vega Notional}) \times \frac{\sigma_{0,T}^2 - K_{var}^2}{2K_{var}} \times 100$$

We will see in subsection 3.2.2 the reasons of this choice.

3.2.1 Replication and fair strike of a Variance Swap

Assume that Call and Put options on the underlying are available in the market. For the moment, let us assume the unrealistic hypothesis that Calls are available for every strike $K \geq S^*$ and Puts are available for every strike $K \leq S^*$ (where S^* is likely to be the ATM or ATMf).

Theorem 3.5. *The realized variance $\sigma_{0,T}^2$ can be replicated with the following instruments: a zero-coupon bond, cash, a forward, a delta strategy, vanilla Put and vanilla Call options on the underlying. The replication is given by:*

$$\boxed{\begin{aligned}\sigma_{0,T}^2 = \frac{2}{T}\Bigg[\int_0^T \frac{dS_t}{S_t} - \ln\frac{S^*}{S_0} - \frac{S_T - S^*}{S^*} + \\ + \int_0^{S^*} \frac{1}{K^2}(K - S_T)_+ dK + \int_{S^*}^{\infty} \frac{1}{K^2}(S_T - K)_+ dK\Bigg]\end{aligned}} \tag{3.7}$$

The previous theorem allows us to derive the fair strike of a Variance Swap. Assume the existence of a constant, risk-free rate r to remain in a simple framework (even though it would be more appropriate to use a stochastic rate), and let Q be the risk-neutral measure. The fair strike K_{var} is the strike which makes the value of the contract equal to zero. The value of the contract is $e^{-rt}E^Q\left[N(\sigma_{0,T}^2 - K_{var}^2)\right]$, therefore the fair strike satisfies

$$K_{var}^2 = E^Q\left[\sigma_{0,T}^2\right] \tag{3.8}$$

Under the risk-neutral measure Q we have that

$$\frac{dS_t}{S_t} = rdt + \sigma_t d\tilde{B}_t$$

where $\tilde{B}_t$ is a Q-Brownian Motion. Therefore,

$$E^Q\left[\int_0^T \frac{dS_t}{S_t}\right] = rT; \tag{3.9}$$

$$E^Q\left[\frac{S_T - S^*}{S^*}\right] = \frac{S_0 e^{rT} - S^*}{S^*}; \tag{3.10}$$

$$E^Q\left[\int_0^{S^*} \frac{1}{K^2}(K - S_T)_+ dK\right] = e^{rT}\int_0^{S^*} \frac{1}{K^2} Put(K)dK \tag{3.11}$$

$$E^Q\left[\int_{S^*}^{\infty} \frac{1}{K^2}(S_T - K)_+ dK\right] = e^{rT}\int_{S^*}^{\infty} \frac{1}{K^2} Call(K)dK \tag{3.12}$$

Combining the previous equations, we find that

$$\boxed{\begin{aligned} K_{var}^2 = \frac{2}{T}\Bigg[&\ln\frac{S_0 e^{rT}}{S^*} - \left(\frac{S_0 e^{rT}}{S^*} - 1\right) + \\ &+ e^{rT}\int_0^{S^*} \frac{1}{K^2} Put(K)dK + e^{rT}\int_{S^*}^{\infty} \frac{1}{K^2} Call(K)dK\Bigg] \end{aligned}} \tag{3.13}$$

3.2.2 Sensitivities

What is in the interest of the trader are the Greeks of the portfolio of Vanilla Options used for replicating the Variance Swap. The portfolio in consideration is

$$\Pi_t = \frac{2}{T}\left[\int_0^{S^*} \frac{1}{K^2} Put_t(K)dK + \int_{S^*}^{\infty} \frac{1}{K^2} Call_t(K)dK\right] \tag{3.14}$$

What the trader will do is to hedge with a discrete version of equation 3.7: holding as many options as possible, suitably weighted[1], and replicating the term $\frac{2}{T}\int_0^T \frac{dS_t}{S_t}$ with the daily rebalanced strategy $\frac{2}{T}\sum_{t=1}^{T} \frac{1}{S_{t-1}}(S_t - S_{t-1})$.

Observation 3.6. The term $\frac{2}{T}\int_0^T \frac{dS_t}{S_t}$ present in the replication of the Variance Swap has different sensitivities than the effective strategy implemented by the trader $\frac{2}{T}\sum_{t=1}^{T} \frac{1}{S_{t-1}}(S_t - S_{t-1})$. For example, the former has a Gamma of $\frac{\partial}{\partial S_t}\left(\frac{2}{TS_t}\right) = -\frac{2}{TS_t^2}$ and the latter has a Gamma of zero.

[1] See section 3.4 for more precise details

Assume for simplicity that interest rates are zero. We can consider equation (3.13) for the period $[t, T]$, obtaining

$$K_{t,T}^2 = \frac{2}{T-t}\left[\ln\frac{S_t}{S^*} - \left(\frac{S_t}{S^*} - 1\right) + \right.$$
$$\left. + \int_0^{S^*} \frac{1}{K^2} Put_t(K)dK + \int_{S^*}^{\infty} \frac{1}{K^2} Call_t(K)dK\right] \tag{3.15}$$

Hence we find

$$\Pi_t = \frac{T-t}{T} K_{t,T}^2 - \frac{2}{T}\ln\frac{S_t}{S^*} + \frac{2}{T}\left(\frac{S_t}{S^*} - 1\right) \tag{3.16}$$

From the above equation we can find all the Greeks.

Vega

The Vega of a Variance Swap coincides with the Vega of the option portfolio, because in the replication (3.7) it is the only vega-sensitive term.

$$\nu_t = \frac{\partial \Pi}{\partial K_{t,T}} = 2\frac{T-t}{T} K_{t,T} \tag{3.17}$$

Very often the notional of Variance Swaps is expressed as a Vega Notional divided by 2 times the Variance Swap initial strike. To clarify,

$$\text{Notional} = 100\frac{\text{Vega Notional}}{2K_{0,T}}$$

This can be explained looking at the above sensitivity: at inception, the Vega of $100\frac{N_{Vega}}{2K}$ Variance Swaps is exactly $100N_{vega}$. Therefore, if volatility increases by one point, the value of the contract increases by $100N_{vega} \times 1\% = N_{vega}$. This means that the Vega Notional is the amount of money which is gained or lost when volatility moves by 1%. However this is only true at inception and not anymore during the life of the contract.
Another way to derive the Vega of a Variance Swap is by using the additivity property (3.2). Since $\sigma_{0,T}^2 = \frac{t}{T}\sigma_{0,t}^2 + \frac{T-t}{T}\sigma_{t,T}^2$, the Mark-to-Market at time t of the variance is

$$MTM_t = \frac{t}{T}\sigma_{0,t}^2 + \frac{T-t}{T} K_{t,T}^2 \tag{3.18}$$

Therefore the Vega of the Variance Swap is

$$\nu_t = \frac{\partial MTM_t}{\partial K_{t,T}} = 2\frac{T-t}{T} K_{t,T} \tag{3.19}$$

Delta

$$\Delta_t = \frac{\partial \Pi}{\partial S_t} = -\frac{2}{TS_t} + \frac{2}{TS^*} \tag{3.20}$$

We can observe that the Delta of the option portfolio is hedged by the delta strategy $\int \frac{dS_t}{S_t}$ and the forward present in the Variance Swap replication (3.7). After all, the Delta of the Variance Swap must be zero.

Gamma

$$\Gamma_t = \frac{\partial^2 \Pi}{\partial S_t^2} = \frac{2}{TS_t^2} \tag{3.21}$$

Gamma is always positive, but it will be balanced by the always negative Theta. We observe that the Gamma faced by the trader is not zero, whereas **the Gamma of a Variance Swap is zero** (having a Variance Swap a constantly zero Delta). The discrepancy is caused by the fact that the trader implements a daily rebalanced strategy $\frac{2}{T}\sum_{t=1}^{T} \frac{1}{S_{t-1}}(S_t - S_{t-1})$ which has zero Gamma. If the trader implemented the continuously rebalanced strategy $\frac{2}{T}\int_0^T \frac{dS_t}{S_t}$ he would see an additional Gamma of $\frac{\partial}{\partial S_t}\left(\frac{2}{TS_t}\right) = -\frac{2}{TS_t^2}$ which would bring the total Gamma to zero.

Theta

$$\theta_t = \frac{\partial \Pi}{\partial t} = -\frac{1}{T}K_{t,T}^2 \tag{3.22}$$

We see that Theta is always negative, and we see that the Gamma profit $\frac{1}{2}\Gamma S^2 \sigma^2 dt$ is equal to the Theta loss $-\theta dt$.
Although less intuitively, the Theta can also be found thanks to the additivity property. From equation (3.18), we have

$$MTM_{t+dt} = \frac{t+dt}{T}\sigma_{0,t+dt}^2 + \frac{T-t-dt}{T}K_{t+dt,T}^2 \tag{3.23}$$

By another application of the additivity property, we have

$$(t+dt)\sigma_{0,t+dt}^2 = t\sigma_{0,t}^2 + dt \cdot \sigma_{t,t+dt}^2 \tag{3.24}$$

When computing the Theta, we assume that spot and volatility do not move, i.e. $\sigma_{t,t+dt}^2 = 0$ and $K_{t+dt,T}^2 = K_{t,T}^2$. Hence

$$\theta_t = \frac{MTM_{t+dt} - MTM_t}{dt} = -\frac{1}{T}K_{t,T}^2 \tag{3.25}$$

Vanna

The Vanna of the option portfolio coincides with the Vanna of the Variance Swap.

$$Vanna_t = \frac{\partial^2 \Pi}{\partial S_t \partial K_{t,T}} = 0 \tag{3.26}$$

This means that the Vega of the Variance Swap is independent of the spot, which translates the fact that a Variance Swap give pure exposure to volatility.

3.2.3 Derivation of the replication in a stochastic volatility model

Assume for the Underlying price a continuous diffusion model with no jumps, with the volatility being a stochastic process. That is, we assume that the price evolution is given by

$$\frac{dS_t}{S_t} = \mu_t dt + \sigma_t dB_t$$

where B_t is a standard Brownian Motion under the Historical Probability.

The realized variance is the sum of a delta strategy and a log contract

By Ito's formula we have

$$d\ln(S_t) = (\mu_t - \frac{1}{2}\sigma_t^2)dt + \sigma_t dW_t = \frac{dS_t}{S_t} - \frac{1}{2}\sigma_t^2 dt \tag{3.27}$$

By integrating both sides of the previous equation from t to T we obtain

$$\ln\frac{S_T}{S_0} = \int_0^T \frac{dS_t}{S_t} - \frac{1}{2}\int_0^T \sigma_t^2 dt \tag{3.28}$$

The realized variance over the period $[0,T]$ is $\sigma_{0,T}^2 = \frac{1}{T}\int_0^T \sigma_t^2 dt$, therefore we obtain that

$$\sigma_{0,T}^2 = \frac{2}{T}\left[\int_0^T \frac{dS_t}{S_t} - \ln\frac{S_T}{S_0}\right] \tag{3.29}$$

The second term in the brackets is an European option on the underlying, since it is a function of the final price of S. We will refer to it as a *log contract*.
The first term in the brackets can be thought as the final value of a Delta

strategy in the underlying. This strategy has initial value of rTe^{-rT}, rebalances continuously and is instantaneously long $\frac{1}{S_t}e^{-r(T-t)}$ shares of the underlying. In fact, let V_t be the value of a self-financed portfolio holding $\frac{1}{S_t}e^{-r(T-t)}$ shares and $V_0 = rTe^{-rT}$; being the strategy self-financed, we can write

$$d(e^{-rt}V_t) = \frac{1}{S_t}e^{-r(T-t)}d(e^{-rt}S_t) = \frac{1}{S_t}e^{-r(T-t)}(e^{-rt}dS_t - re^{-rt}S_t dt) \tag{3.30}$$

By integrating both sides we obtain

$$V_T e^{-rT} - V_0 = e^{-rT}\int_0^T \frac{dS_t}{S_t} - rTe^{-rT} \tag{3.31}$$

and finally, substituting V_0, we find

$$V_T = \int_0^T \frac{dS_t}{S_t}$$

The log contract can be replicated with a strip of vanilla Calls and Puts

The log contract is an European payoff, therefore we can replicate it with vanilla options according to Carr's formula (2.1). $G(x) = -\ln\frac{x}{S_0}$ is twice differentiable, with $G'(x) = -\frac{1}{x}$ and $G''(x) = \frac{1}{x^2}$. Hence,

$$\begin{aligned} -\ln\frac{S_T}{S_0} &= -\ln\frac{S^*}{S_0} - \frac{S_T - S^*}{S^*} \\ &\quad + \int_0^{S^*}\frac{1}{K^2}(K - S_T)_+ dK + \int_{S^*}^{\infty}\frac{1}{K^2}(S_T - K)_+ dK \end{aligned} \tag{3.32}$$

Combining this result with equation (3.29) concludes the proof.

3.3 Variations of Variance Swaps

In this section we will show results similar to the already obtained on Variance Swaps, but applied to variations such as *Conditional Variance Swaps* and *Gamma Swaps*.

3.3.1 Corridor Variance Swaps Definitions

Corridor variance swaps make it possible to be exposed to the variance conditionally to the price of the underlying being in a specified range. For example, an investor might want to sell realised variance, but to freeze his exposure in the event of a market crash. He could achieve this by entering an *Up Variance Swap*, which will provide exposure to the "normal conditions" volatility and not to the "stress conditions" volatility.

Definition 3.7. The *(non-normalized)* corridor realised variance of an underlying with price S_t, with barriers L_{down}, L_{up} is the following quantity:

$$\sigma^2_{non-norm} = \frac{252}{T-t} \sum_{u=t+1}^{T} \ln\left(\frac{S_u}{S_{u-1}}\right)^2 1_{\{L_{down} \leq S_{u-1} \leq L_{up}\}} \tag{3.33}$$

The *(non-normalized)* **Up Variance** is a corridor variance with $L_{up} = +\infty$ and the *(non-normalized)* **Down Variance** is a corridor variance with $L_{down} = 0$.

Definition 3.8. A *(non-normalized)* Corridor Variance Swap is a forward contract on the realised Corridor Variance. Hence the payoff of an Corridor Variance Swap with maturity T, strike K^2_{corr}, notional N and barriers L_{down}, L_{up} is

$$N\left(\sigma^2_{non-norm} - K^2_{corr}\right) \tag{3.34}$$

A Down Variance Swap and an Up Variance Swap are, similarly, forward contracts on the Down Variance and the Up Variance.

In order to have a quantity which is comparable to realised variance, it is more meaningful to normalise by taking a conditional average of the squared log returns:

Definition 3.9. The Corridor Realised Variance of an underlying with price S_t, with barriers L_{down}, L_{up} is the following quantity:

$$\sigma^2_{corr} = \frac{252}{\sum\limits_{u=t+1}^{T} 1_{\{L_{down} \leq S_{u-1} \leq L_{up}\}}} \sum_{u=t+1}^{T} \ln\left(\frac{S_u}{S_{u-1}}\right)^2 1_{\{L_{down} \leq S_{u-1} \leq L_{up}\}} \tag{3.35}$$

The **Up Variance** is a corridor variance with $L_{up} = +\infty$ and the **Down Variance** is a corridor variance with $L_{down} = 0$.

Set

$$N_{cond} = \sum_{u=t+1}^{T} 1_{\{L_{down} \leq S_{u-1} \leq L_{up}\}} \tag{3.36}$$

N_{cond} is the number of days where the underlying lies in the specified range. It can be seen as a strip of digital options, and we have

$$N_{cond}\sigma^2_{corr} = (T-t)\sigma^2_{non-norm}$$

Definition 3.10. A Corridor Variance Swap on the period $[0, T]$, with notional N, has the following payoff:

$$\text{Payoff} = N \cdot \frac{N_{cond}}{T}\left(\sigma^2_{corr} - K^2_{cc}\right) = N\left(\sigma^2_{non-norm} - \frac{N_{cond}}{T}K^2_{cc}\right) \tag{3.37}$$

It is crucial that in the payoff usually there is the factor N_{cond}. This way the Corridor Variance Swap is equal to a non-normalised Corridor Variance Swap plus a strip of digital options. The key problem in both definitions is to replicate the quantity $\sigma^2_{non-norm}$.

3.3.2 Up Variance Replication

The replication of the Up Variance can be obtained through similar mathematical proofs as in the case of standard Variance Swaps. However, this case requires more sophisticated tools in order to deal with the discontinuity at the barrier.
The main ideas in the Variance Swap replication are:

- Apply Ito's formula to the function $G(x) = \ln x$
- Identify the estimator of the variance (the real underlying of the contract) with the theoretical quantity in the model
- Express the realised variance in terms of a delta strategy, plus $G(S_T)$
- Replicate the European Payoff $G(S_T)$ with vanilla options according to Carr's formula

We can use the same ideas, but we need to find a suitable function G. Let L be the down barrier. Intuitively, G must behave like $\ln x$ for $x \geq L$ and $G(x) = 0$ for $x < L$. Most importantly, we ask G to be as regular as possible, in order to apply Ito and Carr formulas. Searching G of the form $G(x) = (a + bx + \ln x)1_{\{x \geq L\}}$, and imposing the constraints that both G and G' be continuous at L, we find

$$G(x) = \left(-\frac{1}{L}(x - L) + \ln\frac{x}{L}\right)1_{\{x \geq L\}}$$

$$G'(x) = \left(-\frac{1}{L} + \frac{1}{x}\right)1_{\{x \geq L\}}$$

G is a concave function; its second derivative in the distributional sense is minus a positive Radon measure, which is absolutely continuous with respect to Lebesgue measure. The Radon-Nikodym derivative is

$$G"(x) = -\frac{1}{x^2} 1_{\{x \geq L\}}$$

We can apply Ito-Tanaka's formula combined with the Occupation Times formula (Equation (2.6)). We obtain:

$$G(S_T) - G(S_0) = \int_0^T G'(S_t) dS_t + \frac{1}{2} \int_0^T G"(S_t) S_t^2 \sigma_t^2 dt$$

$$G(S_T) - G(S_0) = \int_0^T \left(-\frac{1}{L} + \frac{1}{S_t} \right) 1_{\{S_t \geq L\}} dS_t - \frac{1}{2} \int_0^T \sigma_t^2 1_{\{S_t \geq L\}} dt \quad (3.38)$$

In the last term we recognise the Up Variance. The central term is a delta strategy in the underlying. Therefore,

$$\begin{aligned} UpVariance &= \frac{1}{T} \int_0^T \sigma_t^2 1_{\{S_t \geq L\}} dt = \\ &= \frac{2}{T} \left[-G(S_T) + G(S_0) + \int_0^T \left(-\frac{1}{L} + \frac{1}{S_t} \right) 1_{\{S_t \geq L\}} dS_t \right] \end{aligned} \quad (3.39)$$

The last step for finding the replication of the Up Variance with vanilla instruments is to decompose the European Payoff $G(S_T)$ as a combination of Calls and Puts. Since we designed G to be regular enough, we can apply Carr's formula (2.1), with $S^* = L$. Since $G(L) = G'(L) = 0$, we find

$$G(S_T) = -\int_L^{+\infty} \frac{1}{K^2} (S_T - K)_+ dK$$

In conclusion

$$\boxed{UpVar = \frac{2}{T} \left[G(S_0) + \int_0^T \left(-\frac{1}{L} + \frac{1}{S_t} \right) 1_{\{S_t \geq L\}} dS_t + \int_L^{+\infty} \frac{1}{K^2} (S_T - K)_+ dK \right]} \quad (3.40)$$

Observation 3.11. Should there not be any available Call option with strike between L and 100%, it is always possible to build a synthetic Call option using a Put, the Forward and cash, thanks to the Call-Put parity.

3.3.3 Weighted Variance Swap replication

Inspired by the ideas in the previous section, we can generalize the Ito-Tanaka combination with Carr formula methodology.
Take $G : \mathbb{R}^+ \to \mathbb{R}$ satisfying the following hypotheses:

- G is continuous and differentiable
- G is the difference of two convex functions (automatically satisfied if G is C^2);
- the second derivative of G in the distributional sense (which is a Radon measure) is absolutely continuous with respect to Lebesgue measure, with Radon-Nikodym derivative denoted as G".

By equation (2.6) we have that

$$G(S_T) - G(S_0) = \int_0^T G'_-(S_t)dS_t + \frac{1}{2}\int_0^T G"(S_t)S_t^2\sigma_t^2 dt \tag{3.41}$$

Set $\boxed{w(x) := -G"(x)x^2}$. Then we find

$$\frac{1}{T}\int_0^T w(S_t)\sigma_t^2 dt = \frac{2}{T}\left[-G(S_T) + G(S_0) + \int_0^T G'(S_t)dS_t\right] \tag{3.42}$$

As we previously did, we decompose the European Payoff $G(S_T)$ as a combination of Calls and Puts. Since we designed G to be regular enough, we can apply Carr's formula (2.1). We find

$$\begin{aligned} G(S_T) = {} & G(S^*) + G'(S^*)(S_T - S^*) - \\ & - \int_0^{S^*} \frac{w(K)}{K^2}(K - S_T)_+ dK - \int_{S^*}^{\infty} \frac{w(K)}{K^2}(S_T - K)_+ dK \end{aligned} \tag{3.43}$$

Combining the previous two, we finally get

$$\boxed{\begin{aligned} \frac{1}{T}\int_0^T w(S_t)\sigma_t^2 dt = \frac{2}{T}\Bigg[& G(S_0) - G(S^*) - G'(S^*)(S_T - S^*) + \int_0^T G'(S_t)dS_t + \\ & + \int_0^{S^*} \frac{w(K)}{K^2}(K - S_T)_+ dK + \int_{S^*}^{\infty} \frac{w(K)}{K^2}(S_T - K)_+ dK \Bigg] \end{aligned}} \tag{3.44}$$

Definition 3.12. The Weighted Realised Variance of an underlying S_t with weighting $w(x)$ is the following quantity:

$$\sigma_w^2 = \frac{252}{T}\sum_{u=1}^{T} w\left(S_{u-1}\right)\left(\ln\frac{S_u}{S_{u-1}}\right)^2 \tag{3.45}$$

Similarly to what we did in the case of the Variance Swap, we can identify the Weighted Realised Variance with the quantity $\dfrac{1}{T}\displaystyle\int_0^T w(S_t)\sigma_t^2 dt$. The previous equation gives us the way to replicate any Weighted Realised Variance:

1. Identify the weight $w(x)$;
2. Solve the differential equation $G"(x) = -\dfrac{w(x)}{x^2}$ (initial conditions do not matter[2]); by solving we mean finding a function G such that:
   (a) G is continuous and differentiable
   (b) G is the difference of two convex functions (automatically satisfied if G is C^2);
   (c) the second derivative of G in the distributional sense (which is a Radon measure) is absolutely continuous with respect to Lebesgue measure, with Radon-Nikodym derivative equal to $-\dfrac{w(x)}{x^2}$;
3. Use equation (3.44) to find the replication.

3.3.4 Gamma Swaps

Gamma Swaps are an alternative to Variance Swaps for getting exposure to future realized variance. The main advantages of a Gamma Swap is that it provides exposure to realised variance and it is perfectly replicable with Vanilla options (as the Variance Swap), but it avoids explosions in market crashes and usually has a lower strike. The Gamma Swap market has never reached a satisfactory liquidity, however Gamma Swaps are still found in many OTC trades.

Definition 3.13. A Gamma Swap on the underlying S with maturity T is a forward contract on the following quantity[3]:

$$\varsigma^2 = \frac{252}{T}\sum_{t=1}^{T}\frac{S_t}{S_0}\left[\log\frac{S_t}{S_{t-1}}\right]^2 \tag{3.46}$$

[2] By simple calculations, the reader can see that equation (3.44) is invariant for the transformation $G(x) \to \tilde{G}(x) = G(x)+\alpha+\beta x$. Therefore, any G satisfying $G"(x) = -\frac{w(x)}{x^2}$ will produce the same result.

[3] Alternative definitions include $\varsigma^2 = \frac{252}{T}\sum_{t=1}^{T}\frac{S_{t-1}}{S_0}\left[\log\frac{S_t}{S_{t-1}}\right]^2$. The difference is minimal.

The above quantity is very similar to the realized variance. The difference is that log-returns are weighted by the relative move in the spot. This way, when the spot goes down and volatility increases, the Gamma Swap reduces its volatility exposure and avoids blowing up like a Variance Swap. The replication of a Gamma Swap is easily obtained from equation (3.44), with a weight $w(x) = \frac{x}{S_0}$. A function G which satisfies the differential equation

$$G\text{''}(x) = -\frac{w(x)}{x^2} = -\frac{1}{S_0 x} \tag{3.47}$$

is easily found, for example $G(x) = -\frac{x}{S_0}\left(\log\frac{x}{S_0} - 1\right)$, which satisfies all the requested hypotheses. In the case of the Gamma Swap, we find

$$\begin{aligned}\frac{1}{T}\int_0^T \frac{S_t}{S_0}\sigma_t^2 dt = \frac{2}{T}\Bigg[&G(S_0) - G(S^*) - G'(S^*)(S_T - S^*) + \int_0^T G'(S_t)dS_t + \\ &+ \frac{1}{S_0}\int_0^{S^*}\frac{1}{K}(K - S_T)_+ dK + \frac{1}{S_0}\int_{S^*}^{\infty}\frac{1}{K}(S_T - K)_+ dK\Bigg]\end{aligned} \tag{3.48}$$

We see that the weights of the Vanillas are $\frac{1}{K}$ compared to the $\frac{1}{K^2}$ of the Variance Swap. This means that the Gamma Swap replication makes use of less OTM Puts and more OTM Calls than the Variance Swap, and so it has less sensitivity to the skew and lower strike.
In the particular case $S^* = S_0$, we have $G'(S_0) = 0$ and from equation (3.48):

$$\boxed{\begin{aligned}&\text{Gamma Swap Quantity} = \varsigma^2_{[0,T]} = \frac{252}{T}\sum_{t=1}^{T}\frac{S_t}{S_0}\left[\log\frac{S_t}{S_{t-1}}\right]^2 \approx \\ &\approx \frac{2}{S_0 T}\left[-\int_0^T \log\frac{S_t}{S_0}dS_t + \int_0^{S_0}\frac{1}{K}(K - S_T)_+ dK + \int_{S_0}^{+\infty}\frac{1}{K}(S_T - K)_+ dK\right]\end{aligned}} \tag{3.49}$$

Weighted additivity property for Gamma Swaps

Similarly to what we proved in Proposition 3.3, we prove a weighted additivity property for the Gamma Swap quantity

$$\varsigma^2_{[t,T]} = \frac{252}{T-t}\sum_{u=t+1}^{T}\frac{S_u}{S_t}\left[\log\frac{S_u}{S_{u-1}}\right]^2$$

Proposition 3.14. ***(Weighted additivity property for a Gamma Swap):*** *The realised Gamma Swap quantity $\varsigma^2_{[t,T]}$ satisfies the following relation:*

$$\varsigma^2_{[0,T]} = \frac{t}{T}\varsigma^2_{[0,t]} + \frac{T-t}{T}\frac{S_t}{S_0}\varsigma^2_{[t,T]} \tag{3.50}$$

Proof. By definition, the right hand side is equal to

$$\frac{t}{T}\cdot\frac{252}{t}\sum_{u=1}^{t}\frac{S_u}{S_0}\ln\left(\frac{S_u}{S_{u-1}}\right)^2 + \frac{S_t}{S_0}\frac{T-t}{T}\cdot\frac{252}{T-t}\sum_{u=t+1}^{T}\frac{S_u}{S_t}\ln\left(\frac{S_u}{S_{u-1}}\right)^2 \tag{3.51}$$

Combining the sums gives exactly the definition of $\varsigma^2_{[0,T]}$. □

Sensitivities

As in the case of a Variance Swap (section 3.2.2), we investigate the Greeks of the Gamma Swap and of option portfolio which replicates the Gamma Swap. The portfolio in consideration is

$$\Pi_t = \frac{2}{S_0 T}\left[\int_0^{S^*}\frac{1}{K}Put_t(K)dK + \int_{S^*}^{\infty}\frac{1}{K}Call_t(K)dK\right] \tag{3.52}$$

Assume for simplicity that interest rates are zero. We can consider equation (3.48) for the period $[t,T]$, with $G(x) = -\frac{x}{S_t}\left(\log\frac{x}{S_t} - 1\right)$ and $S^* = S_0$, obtaining

$$\begin{aligned}\frac{1}{T-t}\int_t^T\frac{S_u}{S_t}\sigma_u^2 du = \frac{2}{T-t}\Bigg[1 - \frac{S_0}{S_t} + \frac{S_T}{S_t}\log\frac{S_0}{S_t} + \int_0^T G'(S_t)dS_t + \\ + \frac{1}{S_t}\int_0^{S_0}\frac{1}{K}(K-S_T)_+dK + \frac{1}{S_t}\int_{S_0}^{\infty}\frac{1}{K}(S_T-K)_+dK\Bigg]\end{aligned} \tag{3.53}$$

Let $K^2_{[t,T]}$ be the strike of the Gamma Swap on the period $[t,T]$: taking the risk-neutral expectation of the above equation, we have

$$K^2_{[t,T]} = \frac{2}{T-t}\left[1 - \frac{S_0}{S_t} + \log\frac{S_0}{S_t}\right] + \frac{T}{T-t}\cdot\frac{S_0}{S_t}\Pi_t \tag{3.54}$$

$$\Pi_t = \frac{S_t}{S_0}\cdot\frac{T-t}{T}K^2_{t,T} - \frac{2}{S_0 T}\left(S_t - S_0 + S_t\log\frac{S_0}{S_t}\right) \tag{3.55}$$

From the above equation we can find all the Greeks. In particular, we find the Delta and the Gamma.

Delta

$$\Delta_t = \frac{\partial \Pi}{\partial S_t} = \frac{1}{S_0} \cdot \frac{T-t}{T} K_{t,T}^2 + \frac{2}{TS_0} \log \frac{S_t}{S_0} \tag{3.56}$$

We observe that the Delta of the option portfolio is not totally hedged by the delta strategy $\int_0^T \log \frac{S_t}{S_0} dS_t$ present in the replication. The Delta of the Gamma Swap is not zero, but $\frac{1}{S_0} \cdot \frac{T-t}{T} K_{t,T}^2$. Alternatively, this can easily be seen from the additivity property of the Gamma Swap.

Gamma

$$\Gamma_t = \frac{\partial^2 \Pi}{\partial S_t^2} = \frac{2}{TS_0 S_t} \tag{3.57}$$

Opposed to the Variance Swap, where the Gamma was proportional to $\frac{1}{S_t^2}$, in the case of the Gamma Swap the Gamma is proportional to $\frac{1}{S_t}$. The *share Gamma* is defined by the rate of change in the Delta for a small *percent* return ($\frac{dS_t}{S_t}$) of the underlying, i.e.

$$\text{Share Gamma} = S_t \Gamma = \frac{\partial \Delta}{\frac{\partial S_t}{S_t}}$$

We see that in the case of the Gamma Swap, the share Gamma is constant.

Observation 3.15. Similarly to the case of the Variance Swap, there is a discrepancy between the Gamma faced by the trader and the Gamma of a Gamma Swap. The Gamma Swap's Delta is equal to

$$\text{Gamma Swap Delta} = \frac{1}{S_0} \cdot \frac{T-t}{T} K_{t,T}^2$$

Hence, **the Gamma of a Gamma Swap is equal to zero**. The trader will replicate the strategy $-\frac{2}{T} \int_0^T \log \frac{S_t}{S_0} dS_t$ with a daily rebalanced strategy $-\frac{2}{T} \sum_{t=1}^{T} \log \frac{S_{t-1}}{S_0} (S_t - S_{t-1})$ which has zero Gamma. If he instead rebalanced continuously, his additional Gamma would be $-\frac{\partial}{\partial S_t} \left(\log \frac{S_t}{S_0} \right) = -\frac{2}{TS_0 S_t}$. Therefore his total Gamma would be zero.

3.3.5 Corridor Variance Swap Replication

In this case the weight is $w(x) = 1_{\{L \leq x \leq U\}}$. The differential equation $G"(x) = -\frac{w(x)}{x^2}$ is solved by

$$G(x) = \begin{cases} 0 & \text{if } 0 \leq x \leq L \\ -\frac{1}{L}(x-L) + \ln\frac{x}{L} & \text{if } L \leq x \leq U \\ \left(\frac{1}{U} - \frac{1}{L}\right)x + \ln\frac{U}{L} & \text{if } U \leq x \end{cases} \tag{3.58}$$

Therefore,

$$\boxed{\begin{aligned} CorrVar &= \frac{1}{T}\int_0^T 1_{\{L \leq S_t \leq U\}} \sigma_t^2 dt = \\ &= \frac{2}{T}\Bigg[G(S_0) - G(S^*) - G'(S^*)(S_T - S^*) + \int_0^T G'(S_t)dS_t + \\ &+ \int_L^{S^*} \frac{1}{K^2}(K - S_T)_+ dK + \int_{S^*}^U \frac{1}{K^2}(S_T - K)_+ dK\Bigg] \end{aligned}} \tag{3.59}$$

3.3.6 Re-finding the Variance Swap

In this case the weight is $w(x) = 1$. The differential equation $G"(x) = -\frac{1}{x^2}$ is solved by $G(x) = \ln x$ and equation (3.44) yields the same result as previously found in equation (3.7).

3.3.7 Volatility Swaps

Definition 3.16. A Volatility Swap is a forward contract on realized volatility, i.e. on the square root of realized variance.

Let K_{vol} be the strike price of the Volatility Swap and $\sigma_{0,T}^2$ the realized variance of the Underlying over the period $[0, T]$. Then the payoff at maturity of a Variance Swap expiring at time T is

$$N(\sqrt{\sigma_{0,T}^2} - K_{vol})$$

where N is the notional of the contract.
The interest of a volatility swap lies in the fact that it reduces the losses of the seller due to a potential spike in realised variance. In the 2008 financial crisis, investment banks incurred in huge losses on their short positions on variance swaps, and later on variance swaps have always been capped. Volatility

swaps are a safer alternative to sell: if volatility increases from 20% to 60%, a volatility swap pays 40% whereas a variance swap with same vega notional pays $\frac{1}{2\times 20\%}((60\%)^2 - (20\%)^2) = 80\%$.

Strike of a Volatility Swap

Volatility swaps are not replicable with vanilla options like variance swaps. A volatility swap can be seen as a derivative on variance as underlying, therefore it is sensitive to the so-called *vol of vol* (or *vol of var*). A model of diffusion of variance is needed for the fair strike of the volatility swap. The Heston model is a classical example of a stochastic volatility model, but even with the model assumptions there is no closed-form formula for the strike of a volatility swap.
What can be said independently of the model is that the fair strike of a volatility swap is always lower than the fair strike of a variance swap. There are two easy ways to see it:

- (Financial Way) A long position on volatility is short vol of vol, therefore the fair strike has to be reduced to compensate the investor for the sale of vol of vol.
- (Mathematical Way) Let σ be the future realised volatility. We have

$$0 \leq Var[\sigma] = E[\sigma^2] - E[\sigma]^2 = K_{var}^2 - K_{vol}^2 \tag{3.60}$$

  That proves the inequality. Moreover, we see that the difference between the two squared strikes is the variance of the future realised volatility. The strikes coincide only if σ is deterministic (i.e. zero vol of vol, but also no point in trading variance swaps). The more vol of vol, the higher $Var[\sigma]$, the higher the spread.

What we just proved can be put in a proposition:

Proposition 3.17. *The* ***convexity bias****, i.e. the spread between the strikes of a Variance Swap and a Volatility Swap on the same underlying and same maturity, is*

$$Convexity\ Bias = K_{var}^2 - K_{vol}^2 = Var[\sigma_{0,T}] \tag{3.61}$$

where the Variance of σ is taken under the risk-neutral measure.

3.4 Backtest

We backtested the replication of realized variance explained in Theorem 3.5. The strip of options is approximated by a discrete number of options weighted by $\frac{1}{K^2}$. This is the most crucial approximation in the replication, and the results will vary dramatically with the choice of the discretization.

Details of the Backtest

- The underlying is the EuroStoxx 50 Index and the S&P 500 Index (2 backtests).

- The realised variance is the 6-month realised variance, computed with daily close-to-close log returns.

- The term $\int_0^T \frac{dS_t}{S_t}$ is approximated with a daily rebalanced strategy, that is
$$\int_0^T \frac{dS_t}{S_t} \approx \sum_{t=1}^{T} \frac{S_t - S_{t-1}}{S_{t-1}} \tag{3.62}$$

- S^* is the ATM strike.

- The replication with $2n$ options and step h consists in holding n Puts and n Calls. The strikes of the Puts are
$$K_n^p = 1-(n-1)h, \quad ... , \quad K_2^p = 1-h, \quad K_1^p = 1$$
with respective weights $\dfrac{h}{(K_n^p)^2}, ..., \dfrac{h}{(K_2^p)^2}, \dfrac{h}{2\,(K_1^p)^2}$. The strikes of the Calls are
$$K_n^c = 1+(n-1)h\ , \quad ... , \quad K_2^c = 1+h, \quad K_1^c = 1$$
with respective weights $\dfrac{h}{(K_n^c)^2}, ..., \dfrac{h}{(K_2^c)^2}, \dfrac{h}{2\,(K_1^c)^2}$. Note that the ATM Call and Put have an additional $\frac{1}{2}$ factor in the weight: this can be explained mathematically by the trapezoid numerical integration formula, or financially by the fact that we have 2 ATM options which buy ATM volatility and we don't want to overweight this ATM volatility. Observe that in any case we are dropping the extremal intervals of the strikes: the strike range of the options in the market is limited and will never cover the $(0, +\infty)$ theoretical integration interval.

3.4.1 Replication with 4 options, step 10%

This replication makes use of 90-100 Puts and 100-110 Calls. Only four options are not enough to replicate the realized variance. The replication is particularly faulty when the stock price falls, as it misses the Put leg with strikes $< 90\%$. The $\frac{1}{K^2}$ weighting gives much weight to far Out-Of-The-Money Puts, therefore when the stock price falls deeply, the replication error is huge. For example, in the 2008 crisis, the replication error has been around 20 variance points for both the SX5E and the SPX.

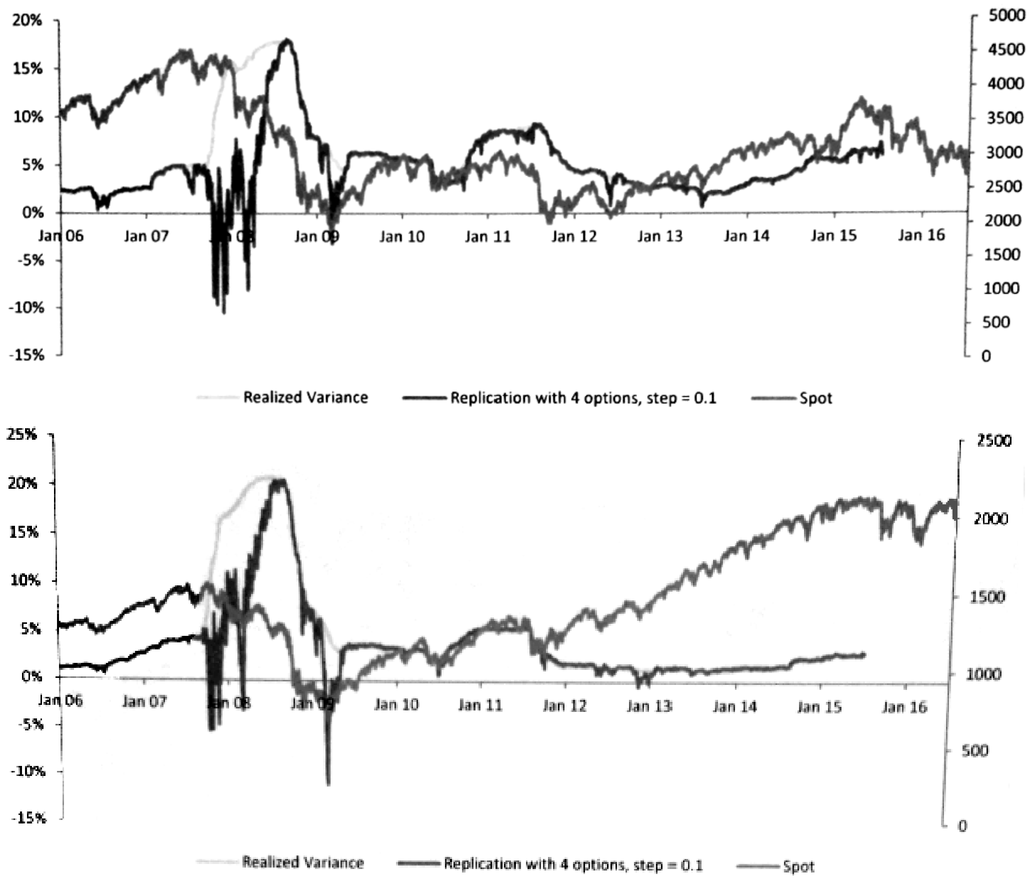


Figure 3.1: 6-month Variance Swap replication on EuroStoxx 50 (top) and S&P 500 (bottom) with 90-100 Puts and 100-110 Calls.

3.4.2 Replication with 8 options, step 5%

This replication makes use of 85-90-95-100 Puts and 100-105-110-115 Calls. Having the 85 Put, the replication is not too bad when the underlying price yearly loss is not more than 15%. However, during the 2008 crisis, yearly losses have been far above 15% and the replication did not work anymore.

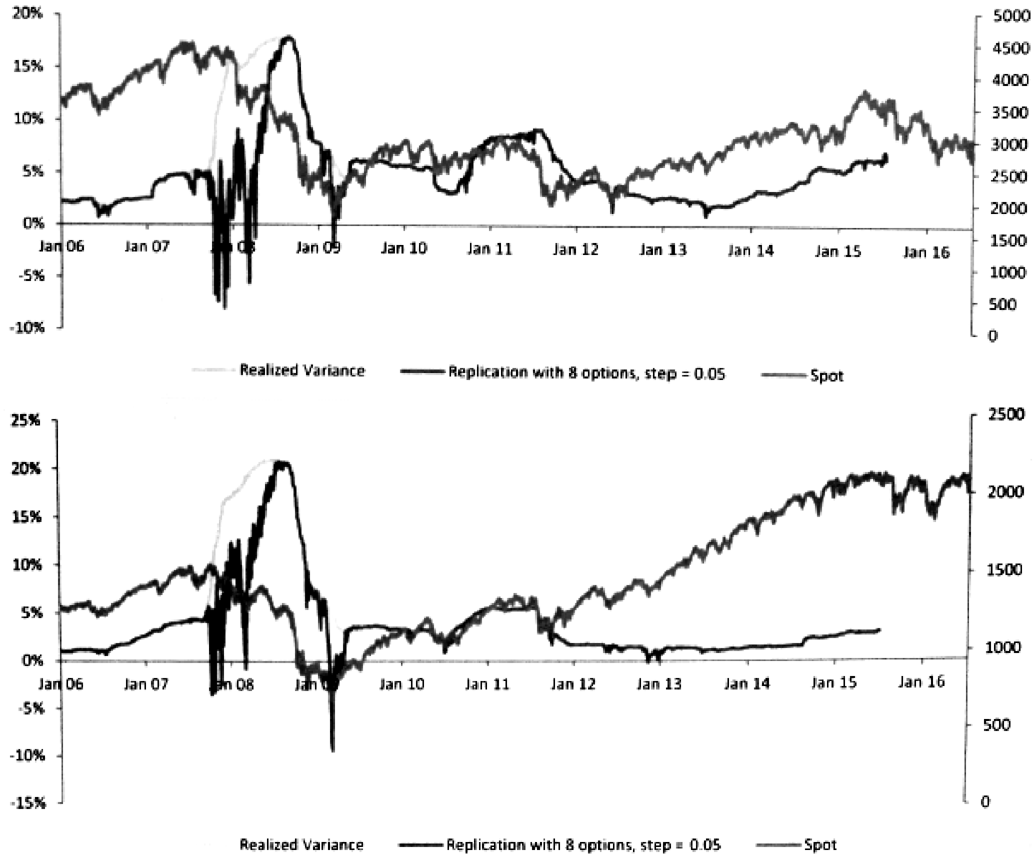


Figure 3.2: 6-month Variance Swap replication on EuroStoxx 50 (top) and S&P 500 (bottom) with 85-90-95-100 Puts and 100-105-110-115 Calls.

3.4.3 Replication with 6 options, step 25%

This replication makes use of 50-75-100 Puts and 100-125-150 Calls. Having the 50 Put, the replication is not impacted during the crisis. However, most of the time the underlying yearly return is in the 80-120 area, where the discretization is too granular. In fact, the replication error is always quite big.

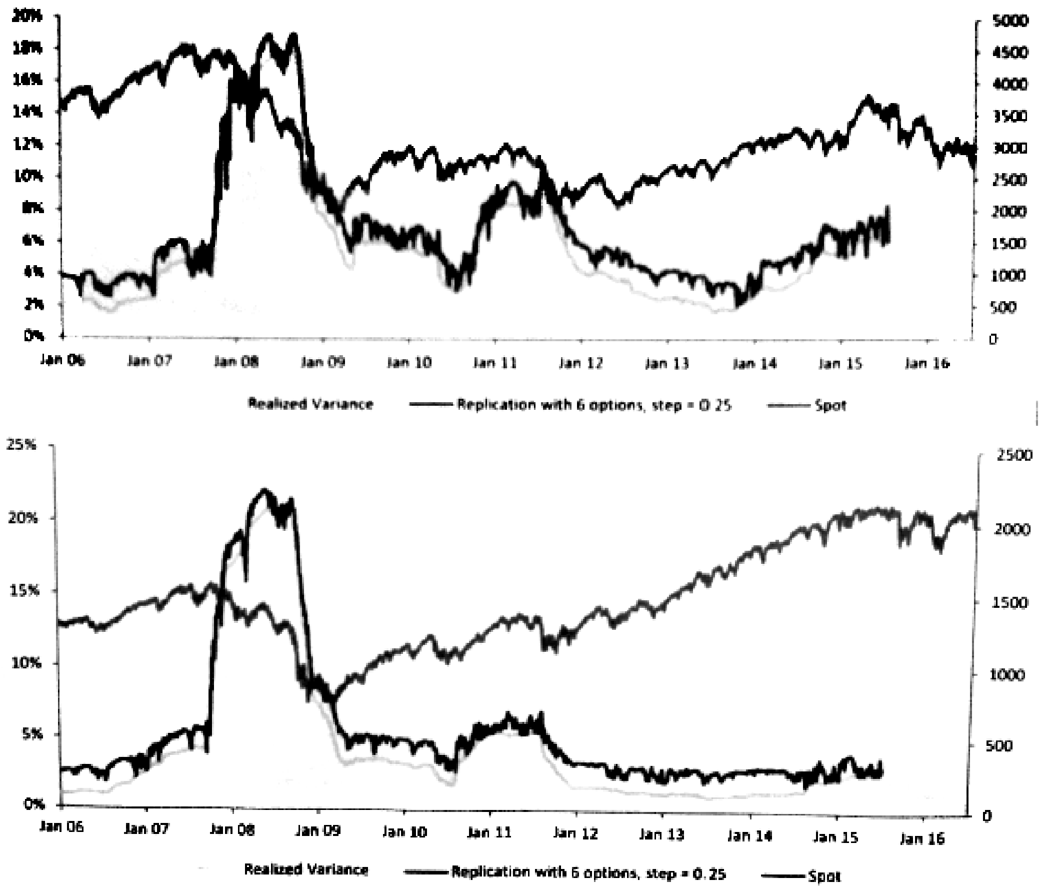


Figure 3.3: 6-month Variance Swap replication on EuroStoxx 50 (top) and S&P 500 (bottom) with 50-75-100 Puts and 100-125-150 Calls.

3.4.4 Replication with 8 options, step 10%

This replication makes use of 70-80-90-100 Puts and 100-110-120-130 Calls. Having the 70 Put, the replication is rarely a disaster: only when the yearly loss is above 30%. Also, the discretization is not too granular near the ATM and the replication error is quite small.

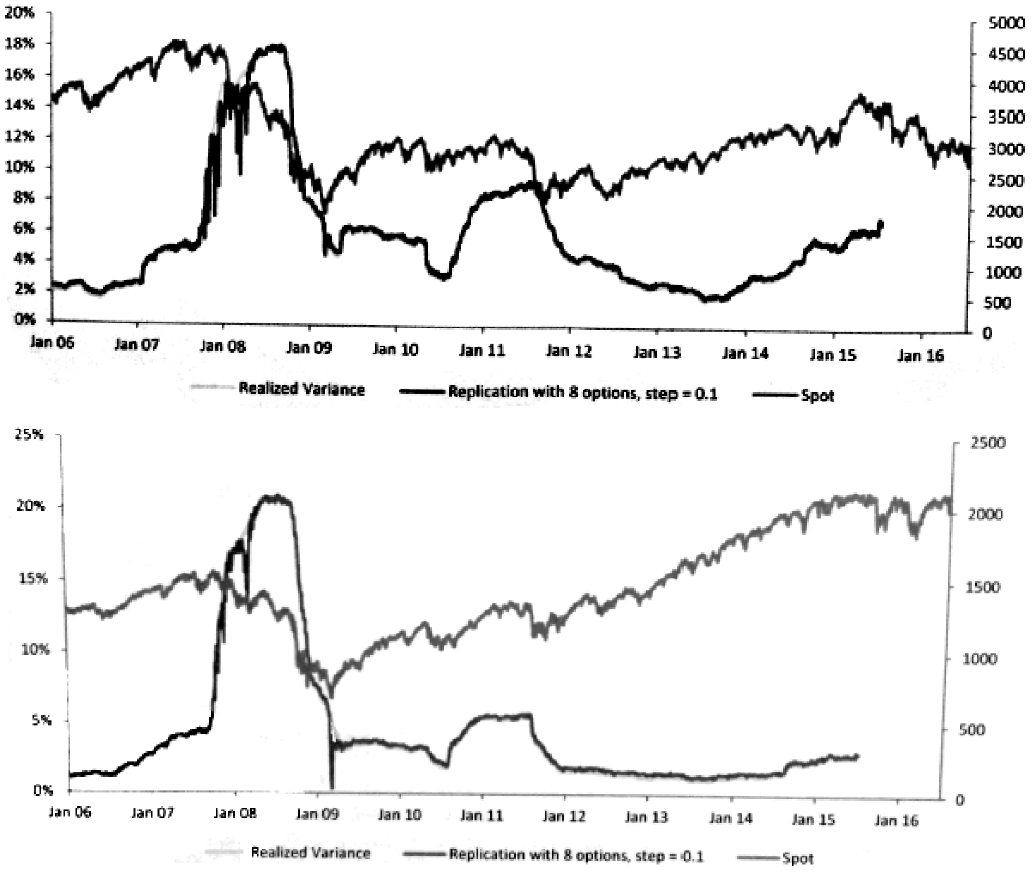


Figure 3.4: 6-month Variance Swap replication on EuroStoxx 50 (top) and S&P 500 (bottom) with 70-80-90-100 Puts and 100-110-120-130 Calls.

3.4.5 Replication with 18 options, step 5%

This replication makes use of 60-65-70-75-80-85-90-95-100 Puts and 100-105-110-115-120-125-130-135-140 Calls. The strike range 60-140 is large enough to cover almost all the realised 1-year movements in the underlying price. In the few cases of larger movements, the far OTM Puts with their large weights help reducing the replication error. The fine discretization in the area near the ATM makes the replication error very small in normal market conditions.

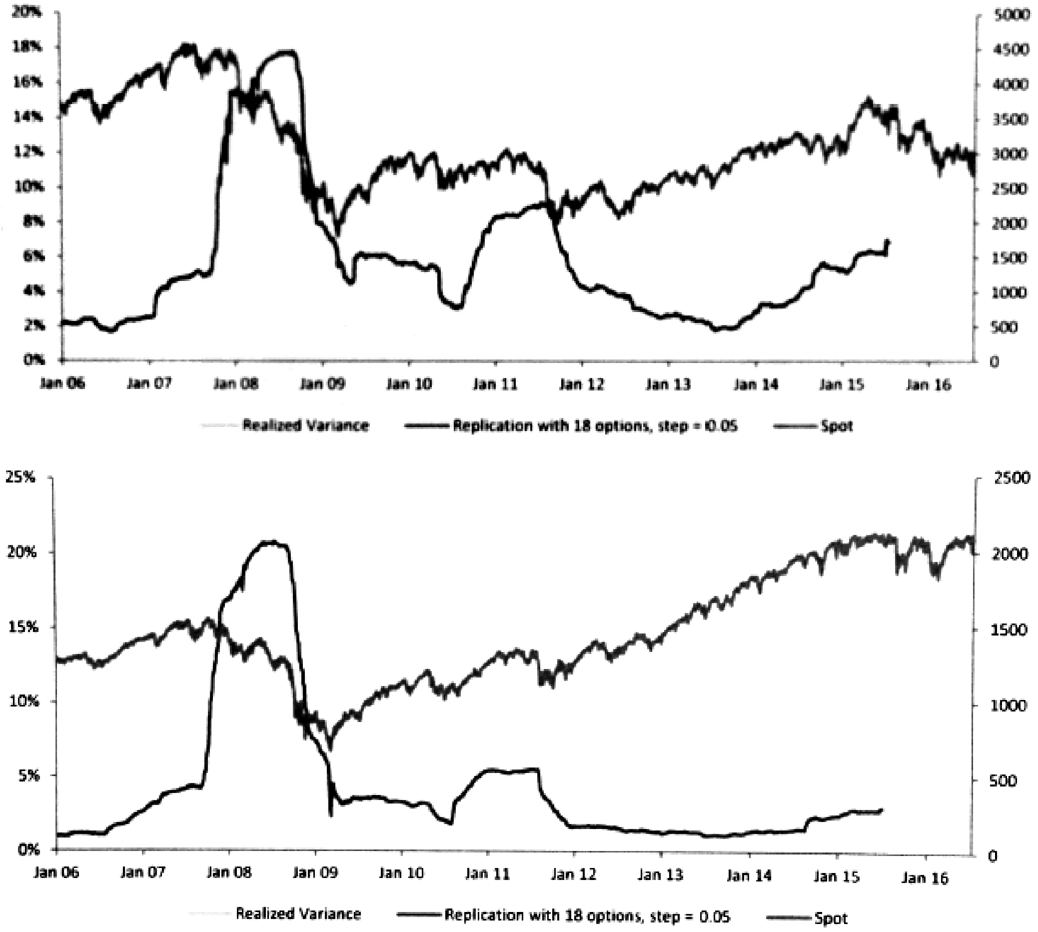


Figure 3.5: 6-month Variance Swap replication on EuroStoxx 50 (top) and S&P 500 (bottom) with 60-65-70-75-80-85-90-95-100 Puts and 100-105-110-115-120-125-130-135-140 Calls.

Chapter 4

Correlation

The *correlation* between the returns of financial assets plays a key role in modern finance. In financial markets, all securities are dependent from each other. Interdependent securities don't necessarily lie within the same asset class, and the degree of dependence, measured by the correlation coefficient, changes continuously.
For example, the price of crude oil and the stock price of an energy company are usually positively correlated, whereas Equities and Bonds are usually negatively correlated. In fact, when Equity markets crash, investors sell their Equity exposure and buy safer bonds, causing bond prices to rise. In periods of bear Equity markets, the correlation between the EuroStoxx 50 and the Bund is very negative, but it may happen to be positive in other market circumstances.
Understanding correlations between financial assets is important to design diversified strategies, to price derivatives on Equity Indices and also to trade correlation itself. Usually, Exotics Trading desks sell correlation to clients and find themselves very *short correlation*. Historically, correlation realizes 5-10 points below implied correlation, which may drive investors to sell the implied and buy the realized. For these and many other reasons, a market of correlation products is born.

4.1 Different types of correlations

4.1.1 Pearson Correlation

In Mathematics there are many ways to express the interdependence of random variables. The most basic and well-known measure is *Pearson's Linear Correlation*, which expresses the **linear** dependence between two random variables. It is defined as:

Definition 4.1. Let X, Y be two random variables. Pearson's linear correlation between X and Y is defined as

$$\rho = \frac{Cov(X,Y)}{\sqrt{Var[X]Var[Y]}} \tag{4.1}$$

Given statistical data $(x_1, y_1), ..., (x_n, y_n)$, the correlation of the data is the natural estimator of the above quantity. Set $\bar{x} = \frac{1}{n}\sum x_i$ and $\sigma^2(x) = \frac{1}{n}\sum(x_i - \bar{x})^2$ (and similarly for y). Then

$$\rho = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sigma(x)\sigma(y)} \tag{4.2}$$

It is important to understand the faults of this measure:

- A correlation of 0 does not mean that the 2 variables are independent;
- This measure only captures the *linear* dependence between the variables;
- The standard estimator of the correlation is very noisy, and very sensitive to data anomalies;
- It is not suited to measuring the interdependence between 3 or more variables.

A very educative example is the following. Let X be a $N(0,1)$ standard Gaussian variable, and $Y = X^2$. Clearly, Y is completely determined by X. However,

$$Cov(X,Y) = E[XY] - E[X]E[Y] = E[X^3] = 0 \tag{4.3}$$

So $\rho = 0$. Only looking at the value of ρ, we would conclude that there is almost independence between X and Y, whereas there is complete dependence!

There are more sophisticated measures of dependence in Mathematics. Copulas are the most complete way to describe the interdependence between n random variables; *Spearman* and *Kendall* correlations are more robust measures than the linear one. However, despite its faults, the market standard is to use Pearson's correlation, whose advantages are its simplicity and its effectiveness in most of the practical cases.

4.1.2 Kendall and Spearman Correlation

Both Spearman correlation ρ_s and Kendall's rank correlation coefficient, or τ coefficient, are measures of the *rank correlation* between two random variables. The two quantities are sensitive only to the ranks of the data, not to the values, and they are less sensitive to data anomalies. While Pearson's correlation captures linear relationships, Spearman and Kendall correlations assess monotonic relationships (whether linear or not). If there are no repeated data values, a perfect Spearman or Kendall correlation of +1 or -1 occurs when each of the variables is a perfect monotone function of the other.

Spearman Correlation

The Spearman correlation coefficient is defined as the Pearson correlation coefficient between the ranked variables. To be precise, take statistical data $(x_1, y_1), ..., (x_n, y_n)$ and assume that no data value is repeated. We rank the x_i and the y_i obtaining $x_{i_1} > x_{i_2} > ... > x_{i_n}$ and $y_{i_1} > y_{i_2} > ... > y_{i_n}$. The rank is defined as $\text{rk}(x_{i_j}) = j$ and similarly for y. In other words, the rank of a x_i is 1 if it is the largest x observation, 2 if it is the second largest etc.

Definition 4.2. The Spearman Correlation of the data $(x_1, y_1), ..., (x_n, y_n)$ is defined as

$$\rho_s = \frac{\frac{1}{n}\sum_{i=1}^{n}(\text{rk}(x_i) - \overline{\text{rk}(x)})(\text{rk}(y_i) - \overline{\text{rk}(y)})}{\sigma(\text{rk}(x))\sigma(\text{rk}(y))} \tag{4.4}$$

Kendall Correlation

Definition 4.3. Let (X, Y) be a random vector. Take (X', Y') an i.i.d. copy of (X, Y). Kendall's τ correlation between X and Y is defined as

$$\tau = E\left[\text{sgn}((X - X')(Y - Y'))\right] \tag{4.5}$$

where sgn denotes the sign function $\text{sgn}(x) = 1_{\{x>0\}} - 1_{\{x<0\}}$.

In practice, given statistical data $(x_1, y_1), ..., (x_n, y_n)$, the τ coefficient is the natural estimator of the quantity in (4.5), i.e.

$$\begin{aligned} \tau &= \frac{1}{\binom{n}{2}} \sum_{i=1}^{n} \sum_{j=1}^{n} \text{sgn}(x_i - x_j)(y_i - y_j) = \\ &= \frac{2}{n(n-1)} \left[(\text{number of concordant pairs}) - (\text{number of discordant pairs})\right] \end{aligned} \tag{4.6}$$

Relationship with Pearson correlation for Gaussian vectors

In this subsection we will find a relationship between Kendall's τ coefficient and Pearson's ρ correlation in the case of a bidimensional Gaussian random vector. Note that this is a special case; it is not possible in general to derive a relationship between the two correlation measures.
The result is valid for a bidimensional Gaussian vector. For completeness, we remind the definition:

Definition 4.4. (X, Y) is said to be a *bidimensional Gaussian vector*, or to have a *bivariate normal distribution*, if for every $(a, b) \in \mathbb{R}^2$, $aX + bY$ is normally distributed.

Remind that two normally distributed random variables need not be a bidimensional Gaussian vector. For example, $X \sim N(0, 1)$ and

$$Y = \begin{cases} X & if\, |X| < 1 \\ -X & if\, |X| \geq 1 \end{cases} \tag{4.7}$$

are both normally distributed but not a Gaussian vector (because $P(X+Y = 0) \neq 0$).

Theorem 4.5. *Let (X, Y) be a Gaussian vector with ρ being Pearson's correlation and τ Kendall's correlation between X and Y. Then*

$$\rho = \sin\left(\frac{\tau\pi}{2}\right) \tag{4.8}$$

Proof. Taken (X', Y') an independent copy of (X, Y) as in the definition of Kendall's τ, set $Z_1 = X - X', Z_2 = Y - Y'$. Since (X, Y) is a Gaussian vector, then also (Z_1, Z_2) is a Gaussian vector, with $E[Z_1] = E[Z_2] = 0$; $Var[Z_1] = 2Var[X]$, $Var[Z_2] = 2Var[Y]$. We have

$$\begin{aligned} E[Z_1 Z_2] &= E[(X - X')(Y - Y')] = \\ &= E[XY] - E[X]E[Y'] - E[X']E[Y] + E[X'Y'] = \\ &= 2(E[XY] - E[X]E[Y]) \end{aligned} \tag{4.9}$$

So, the correlation between Z_1 and Z_2 is the same ρ as the correlation between X and Y. Being both ρ and τ insensitive to positive affine transformations, we can suppose WLOG that (Z_1, Z_2) have zero mean and unit variance. Since (Z_1, Z_2) is a Gaussian vector, we can write

$$Z_2 = \rho Z_1 + \sqrt{1 - \rho^2} Z_3 \tag{4.10}$$

where $Z_3 \sim N(0,1)$. For this step it is crucial that (Z_1, Z_2) is a Gaussian vector, otherwise Z_3 would not be Gaussian.
Now (Z_1, Z_3) is a standard bivariate Gaussian vector, with Z_1, Z_3 uncorrelated therefore independent. Changing the coordinates to polar coordinates, we set

$$\begin{cases} Z_1 = R\cos\theta \\ Z_3 = R\sin\theta \end{cases} \tag{4.11}$$

where R^2 has a $\chi^2(2)$ (or equivalently, exponential) distribution, θ has a uniform distribution on $[-\frac{\pi}{2}, 3\frac{\pi}{2}]$ and R, θ are independent.
We now compute Kendall's τ. Let $\tau' \in [-1,1]$ such that $\rho = \sin\left(\frac{\tau'\pi}{2}\right)$. Our goal is to show that $\tau = \tau'$.

$$\tau = E\left[\mathrm{sgn}(Z_1 Z_2)\right] = P(Z_1 Z_2 > 0) - P(Z_1 Z_2 < 0) = 2P(Z_1 Z_2 > 0) - 1; \tag{4.12}$$

$$\begin{aligned} P(Z_1 Z_2 > 0) &= P\left[Z_1(\rho Z_1 + \sqrt{1-\rho^2} Z_3) > 0\right] = \\ &= P\left[R^2\cos\theta(\rho\cos\theta + \sqrt{1-\rho^2}\sin\theta) > 0\right] = \\ &= P\left[\cos\theta\left(\sin\left(\frac{\tau'\pi}{2}\right)\cos\theta + \cos\left(\frac{\tau'\pi}{2}\right)\sin\theta\right) > 0\right] = \\ &= P\left[\cos\theta\cdot\sin\left(\theta + \frac{\tau'\pi}{2}\right) > 0\right] \end{aligned} \tag{4.13}$$

Now it is a simple trigonometric function sign exercise, which is solved by θ being either in $[-\frac{\tau'\pi}{2}, \frac{\pi}{2}]$ or in $[\pi - \frac{\tau'\pi}{2}, \frac{3\pi}{2}]$ and giving a probability of

$$P(Z_1 Z_2 > 0) = \frac{1}{2\pi}\left[\left(\frac{\pi}{2} + \frac{\tau'\pi}{2}\right) + \left(\frac{3\pi}{2} - \pi + \frac{\tau'\pi}{2}\right)\right] = \frac{1}{2}(1+\tau') \tag{4.14}$$

Hence,

$$\tau = 2\frac{1}{2}(1+\tau') - 1 = \tau' \tag{4.15}$$

□

Example: spot/div correlation

The chart in figure 4.1 analyses the correlation between the EuroStoxx 50 Net Total Return Index and its December 2015 dividend futures. For long term maturities, it is reasonable to think that dividends will be proportional to the spot level. So we expect a very high correlation between the two assets when the maturity is more than 2 years away (which is reflected in the chart

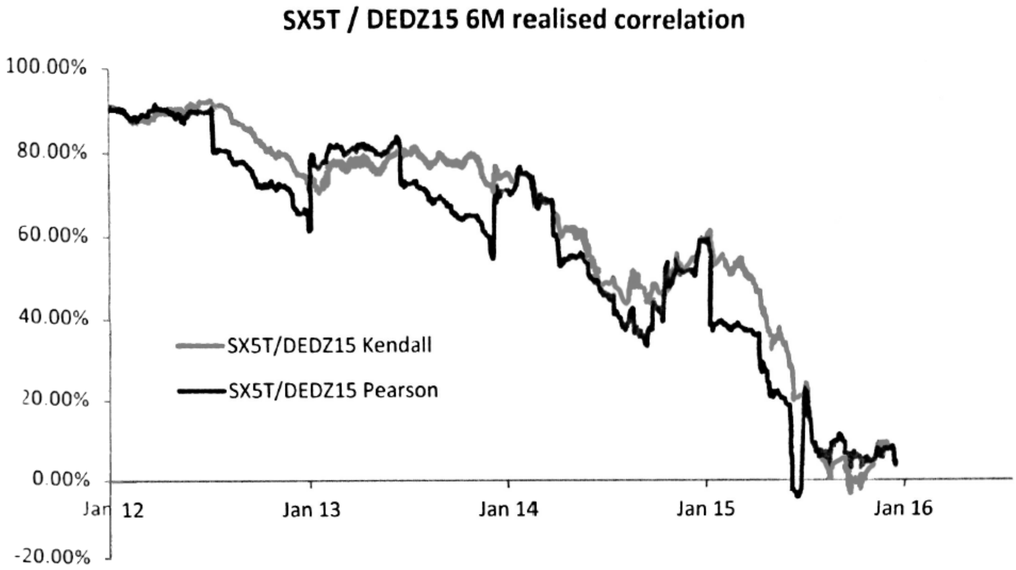


Figure 4.1: 6-month Pearson realized correlation between the EuroStoxx 50 Net Total Return Index (SX5T Index) and the December 2015 EuroStoxx Dividend Futures, compared to the quantity $\sin\left(\frac{\tau\pi}{2}\right)$ where τ is Kendall's correlation coefficient.

in the period preceding January 2014). On the other hand, since incoming dividends are announced before they are paid, for shorter maturities the dividend futures behaves like cash and therefore has zero correlation with the spot price.

Both the previous features are well represented by Pearson and Kendall correlations in the chart. However, on some specific dates, anomalies in the data generate a sharp decrease in the Pearson realised correlation. For example, on July 26, 2012, Telefonica announced that they would stop paying dividends. Dividend futures price sharply dropped, whereas the EuroStoxx had a positive return. Even though it is only one day of data anomaly, it affects significantly the computed Pearson realised correlation for the following 6 months! This is due to the relatively big size of the jump, which overweights the observation at July 26, 2012. Kendall's correlation instead is less sensitive to the data anomaly.

4.1.3 Correlation in the Black-Scholes model

Given two assets, with price processes S_t^1, S_t^2, we can model their evolution with a Black-Scholes model:

$$\begin{cases} \frac{dS_t^1}{S_t^1} = \mu_1 dt + \sigma_1 dW_t^1 \\ \frac{dS_t^2}{S_t^2} = \mu_2 dt + \sigma_2 dW_t^2 \end{cases} \tag{4.16}$$

where W_t^1, W_t^2 are two Brownian motions, not necessarily independent. When saying that S^1 and S^2 have a correlation of ρ, we mean that the joint quadratic variation[1]

$$d[W_t^1, W_t^2] = \rho dt \tag{4.17}$$

At first sight, this may seem unrelated to Pearson correlation. However, if we compute the Pearson correlation of the returns of the assets, assuming that the parameters μ_i, σ_i are constant, we find

$$\begin{aligned} Cov\left(\frac{\Delta S_t^1}{S_t^1}, \frac{\Delta S_t^2}{S_t^2}\right) &= Cov(\sigma_1 \Delta W_t^1, \sigma_2 \Delta W_t^2) = E\left[\sigma_1\sigma_2 \Delta W_t^1 \Delta W_t^2\right] = \\ &= \sigma_1\sigma_2 E\left[\Delta(W_t^1 W_t^2) - W_t^1 \Delta W_t^2 - W_t^2 \Delta W_t^1\right] \end{aligned} \tag{4.18}$$

By Ito's formula, we have that

$$d(W_t^1 W_t^2) = W_t^1 dW_t^2 + W_t^2 dW_t^1 + d[W_t^1, W_t^2]$$
$$\Rightarrow \Delta(W_t^1 W_t^2) \approx W_t^1 \Delta W_t^2 + W_t^2 \Delta W_t^1 + \rho \Delta t$$

Combining the previous equations, we find

$$Cov\left(\frac{\Delta S_t^1}{S_t^1}, \frac{\Delta S_t^2}{S_t^2}\right) = \sigma_1 \sigma_2 \rho \Delta t$$

Since $Var[\frac{\Delta S_t^1}{S_t^1}] = \sigma_1^2 \Delta t$, we conclude that the Pearson correlation of the returns of the assets is exactly the parameter ρ.
Thanks to theorem 4.5, and to the joint Gaussianity of the returns in Black-Scholes model, we can also affirm that

$$\rho = \sin\left(\frac{\tau\pi}{2}\right) \tag{4.19}$$

where τ is Kendall's correlation coefficient between the assets' returns.

Observation 4.6. The previous analysis leads to the same result if we use *log returns* instead of *linear returns*.

[1]See Chapter Four of [1] for the definition of quadratic variation of semimartingales.

4.1.4 Realised correlation and Picking Frequency

Let S_t^1, S_t^2 be the prices of two assets. When we speak about the correlation between S_t^1, S_t^2, we always refer to the correlation of the returns of S_t^1, S_t^2. Given the historical data of the prices of the assets, we can compute their realised correlation. First of all, we need to compute the returns X_t^1, X_t^2 of the assets. In general, it is preferable to use the log returns.

Definition 4.7. Given a time length T (e.g. 6 months), the daily T-realised correlation between S^1 and S^2 at time t is the quantity

$$\rho_t = \frac{\sum_{u=t-T+1}^{t} (X_u^1 - \bar{X^1}_{[t-T+1,t]})(X_u^2 - \bar{X^2}_{[t-T+1,t]})}{\left[\sum_{u=t-T+1}^{t} (X_u^1 - \bar{X^1}_{[t-T+1,t]})^2 \sum_{u=t-T+1}^{t} (X_u^2 - \bar{X^2}_{[t-T+1,t]})^2\right]^{1/2}} \tag{4.20}$$

where

$$X_u^i = \ln \frac{S_u^i}{S_{u-1}^i} \qquad ; \qquad \bar{X^i}_{[t-T+1,t]} = \frac{1}{T} \sum_{u=t-T+1}^{t} X_u^i$$

In fact, the quantity that we are measuring is the theoretical correlation ρ as defined in formula (4.17), through the standard estimator of the correlation.
We can similarly define the realised correlation with Picking Frequency:

Definition 4.8. Given a time length T (e.g. 6 months), the T-realised correlation with Picking Frequency $p \in \mathbb{N}$ between S^1 and S^2 at time t is the quantity

$$\rho_t = \frac{\sum_{u=t-T+p}^{t} (X_{u,p}^1 - \bar{X}^1_{\cdot,p[t-T+p,t]})(X_{u,p}^2 - \bar{X}^2_{\cdot,p[t-T+p,t]})}{\left[\sum_{u=t-T+p}^{t} (X_{u,p}^1 - \bar{X}^1_{\cdot,p[t-T+p,t]})^2 \sum_{u=t-T+p}^{t} (X_{u,p}^2 - \bar{X}^2_{\cdot,p[t-T+p,t]})^2\right]^{1/2}} \tag{4.21}$$

where

$$X_{u,p}^i = \ln \frac{S_u^i}{S_{u-p}^i} \qquad ; \qquad \bar{X}^i_{\cdot,p[t-T+1,t]} = \frac{1}{T} \sum_{u=t-T+p}^{t} X_{u,p}^i$$

This is the same definition as before, except for the fact that instead of *daily* returns, multiple-day returns are used for the estimation. In practice, typical values for the Picking Frequency are between 2 days and 2 weeks.

The Picking Frequency correlation still estimates correlation

The above defined estimator is still measuring the correlation between $X^1_{u,p} = \ln \frac{S^1_u}{S^1_{u-p}}$ and $X^2_{u,p} = \ln \frac{S^2_u}{S^2_{u-p}}$. We have

$$X^i_{u,p} = \ln \frac{S^i_u}{S^i_{u-p}} = \sum_{v=u-p+1}^{u} \ln \frac{S^i_v}{S^i_{v-1}} = \sum_{v=u-p+1}^{u} X^i_v$$

Under the Black-Scholes model, the X^i_v are i.i.d. $N(0, \sigma_i^2/252)$, hence $X^i_{u,p} \sim N(0, \frac{p}{252}\sigma_x^2)$. The covariance between $X^1_{u,p}$ and $X^2_{u,p}$ is

$$\begin{aligned} E\left[X^1_{u,p} X^2_{u,p}\right] &= E\left[\left(\sum_{v=u-p+1}^{u} X^1_v\right)\left(\sum_{v=u-p+1}^{u} X^2_v\right)\right] = \\ &= \sum_{v=u-p+1}^{u} E\left[X^1_v X^2_v\right] = p\rho\sigma_1\sigma_2/252 \end{aligned} \tag{4.22}$$

In the previous steps, mixed terms disappear because returns in different days are independent. Therefore the correlation between $X^1_{u,p}$ and $X^2_{u,p}$ is the same ρ as it is with daily returns.

Why using Picking Frequencies

In the formula (4.20), *daily* log returns are used as a statistical sample for computing the estimator (4.2). Daily returns means specifically close to close returns, i.e. the price used for every day is the price fixing at the close of the exchange where the asset is traded. Using daily returns therefore is appropriate only if the two assets are traded in the same geographical area: we require that the 2 prices at the close are synchronous.
This is not the case when measuring, for example, the correlation between the EuroStoxx 50 and the S&P 500: markets opening hours are far from synchronous in Europe and in the US. As can be seen in figure 4.2, using the daily returns we mistakenly use asynchronous returns in the estimation, thus **underestimating** the true correlation between the assets.

Theorem 4.9. *Suppose that A_t and B_t are the prices of two assets. Assume Black-Scholes dynamics for the two assets, with a correlation of ρ (defined as in section 4.1.3). Let $0 < \alpha < 1$ be a real number, representing a lag between closing hours of the exchanges where A and B are traded, less than 1 day long. Let $\rho_{p,\alpha}$ be the correlation between the close-to-close p-days returns of A and B. To be precise,*

$$\rho_{p,\alpha} = Correl\left(\ln \frac{A_t}{A_{t-p}}, \ln \frac{B_{t+\alpha}}{B_{t-p+\alpha}}\right) \tag{4.23}$$

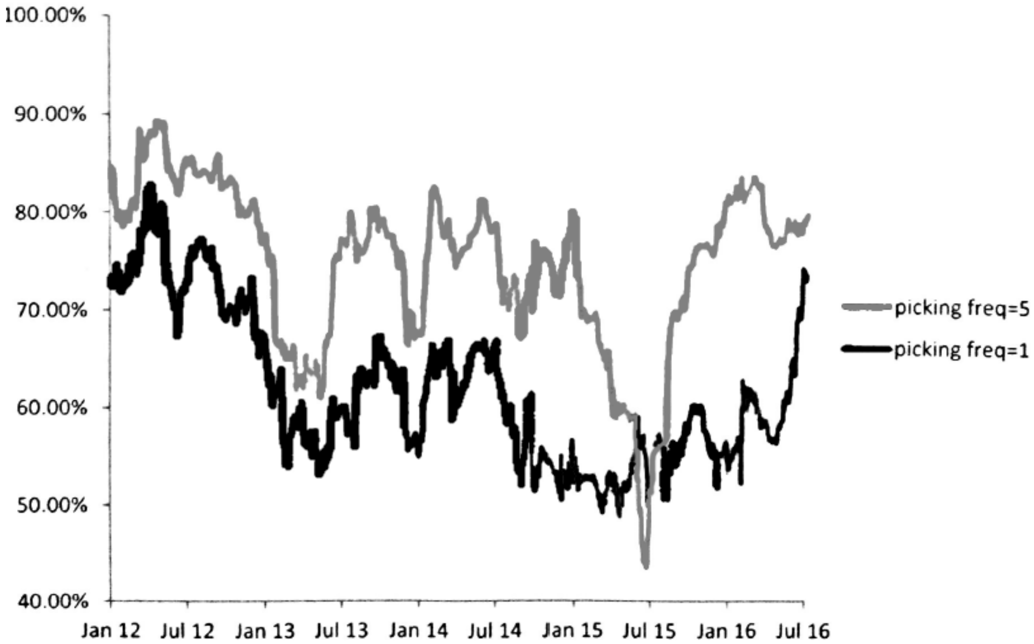


Figure 4.2: 6-month realized correlation between the EuroStoxx 50 and the S&P 500, computed with daily log-returns and weekly log-returns

Then we have

$$\rho_{p,\alpha} = \rho(1 - \frac{\alpha}{p}). \tag{4.24}$$

Proof. We already showed in the previous subsubsection that $\ln \frac{A_t}{A_{t-p}}, \ln \frac{B_{t+\alpha}}{B_{t-p+\alpha}}$ are Gaussian variables with zero mean and variances $\frac{p}{252}\sigma_A^2$, $\frac{p}{252}\sigma_B^2$. The covariance is

$$E\left[\ln \frac{A_t}{A_{t-p}} \ln \frac{B_{t+\alpha}}{B_{t-p+\alpha}}\right] = E\left[\left(\ln \frac{A_t}{A_{t-p+\alpha}} + \ln \frac{A_{t-p+\alpha}}{A_{t-p}}\right)\left(\ln \frac{B_{t+\alpha}}{B_t} + \ln \frac{B_t}{B_{t-p+\alpha}}\right)\right] \tag{4.25}$$

Returns on non-overlapping time periods are always independent, so three terms in the previous equation disappear. The only surviving term is $E\left[\ln \frac{A_t}{A_{t-p+\alpha}} \ln \frac{B_t}{B_{t-p+\alpha}}\right]$. Furthermore, because of what was shown in equation (4.22), substituting $p \to p - \alpha$,

$$E\left[\ln \frac{A_t}{A_{t-p}} \ln \frac{B_{t+\alpha}}{B_{t-p+\alpha}}\right] = E\left[\ln \frac{A_t}{A_{t-p+\alpha}} \ln \frac{B_t}{B_{t-p+\alpha}}\right] = \rho\sigma_A\sigma_B\frac{p-\alpha}{252} \tag{4.26}$$

Finally, we compute

$$\rho_{p,\alpha} = \frac{\rho\sigma_A\sigma_B(p-\alpha)}{\sqrt{p\sigma_A^2 \cdot p\sigma_B^2}} = \rho\left(1 - \frac{\alpha}{p}\right) \tag{4.27}$$

□

Thanks to the previous theorem, we now see why a daily ($p = 1$) correlation underestimates the true correlation when working with the European and American stocks. We also observe that

> The higher the Picking Frequency p, the smaller is the underestimation error made on the correlation.

However, correlation is a variable quantity, so an excessively big Picking Frequency fails to capture the instantaneous correlation. Acceptable values for the Picking Frequency are in the range 3 to 10 days.

4.2 Correlation of $n > 2$ assets

All the definitions for correlation given so far only apply to a *pair* of random variables. We want to generalise this concept for $n > 2$ random variables. In mathematics there is the concept of copulas to describe multi-dimensional dependence between random variables. However, copulas are complicated objects (they are *functions*), whereas we are looking for a single real number to describe the dependence.
It is clear that 3 variables which have pairwise correlations equal to 1 will have a "threewise" correlation equal to 1. However, if the variables are $X_1 = X, X_2 = X, X_3 = -X$, the correlations will be $+1, -1, -1$. Now the meaning of a "threewise" correlation is less clear.
Let us set up the problem within the financial framework: we have $S_t^1, ..., S_t^n$ the prices of n stocks. The daily returns of the stocks are $\frac{\Delta S_t^i}{S_t^i} = \frac{S_{t+1}^i}{S_t^i} - 1$. In what follows, in order for the correlation to have a meaningful sense, we will make the following assumption:

Assumption 4.10. For every pair S^i, S^j of assets, let $\rho_{i,j}$ be the Pearson correlation between the two stocks. From now on, we assume that $\rho_{i,j} \geq 0$ for every i, j and that the values of the $\rho_{i,j}$ are all close to each other (say, no more than 50% difference in any pair of correlations).

This assumption is reasonable when working with equities: usually the considered assets are stocks in the same sector or in the same country and have positive, similar correlations between pairs.

4.2.1 Average Pairwise Correlation

A possible way to define the realised correlation between the n assets is to simply take the average of the $\frac{1}{2}n(n-1)$ pairwise realised correlations. That is

$$\rho = \frac{2}{n(n-1)} \sum_{i \neq j} \rho_{ij} \tag{4.28}$$

The above definition needs assumption 4.10 to make sense. For example, in the case of 3 assets with $S^2 = S^1, S^3 = -S^1$, the above formula would yield a correlation of $-\frac{1}{3}$ which is hard to interpret. But if the correlations were $50\%, 60\%, 55\%$ (these are typical values), a value of 55% for the three-wise correlation is acceptable and has a good interpretation.

4.2.2 Clean and Dirty Correlations

These definitions of correlation are the ones which naturally appear in dispersion strategies, which we will study in the next chapter.
Take $S_t^1, ..., S_t^n$ the prices of n assets. Let B be a dynamic basket of the stocks,as defined in equation (2.16), with **constant weights** $w^1, ...w^n$. Assuming that the daily returns of the assets have a stationary distribution with mean 0 and variances $\sigma_1^2/252, ..., \sigma_n^2/252$, we can compute the volatility of the basket as follows:

$$\begin{aligned} \sigma_B^2 &= 252 Var\left[\frac{\Delta B}{B}\right] = 252 Var\left[\sum_{i=1}^n w^i \frac{\Delta S^i}{S^i}\right] = \\ &= \sum_{i=1}^n (w^i)^2 \sigma_i^2 + 2 \sum_{i<j} w^i w^j \sigma_i \sigma_j \rho_{ij} \end{aligned} \tag{4.29}$$

Assuming that all the correlations are equal, i.e. $\rho_{ij} = \rho$, the above formula becomes

$$\sigma_B^2 = \sum_{i=1}^n (w^i)^2 \sigma_i^2 + 2\rho \sum_{i<j} w^i w^j \sigma_i \sigma_j = \tag{4.30}$$

$$= \sum_{i=1}^n (w^i)^2 \sigma_i^2 + \rho \left[\left(\sum_{i=1}^n w^i \sigma_i \right)^2 - \sum_{i=1}^n (w^i)^2 \sigma_i^2 \right] = \tag{4.31}$$

$$= \rho \left(\sum_{i=1}^{n} w^i \sigma_i \right)^2 + (1-\rho) \sum_{i=1}^{n} (w^i)^2 \sigma_i^2 \tag{4.32}$$

From the above relations we can derive two possible definitions for the n-dimensional correlation. The first one is obtained by deriving ρ from the equation (4.31):

Definition 4.11. The implied/realised *clean correlation* among the n assets is defined as

$$\rho = \frac{\sigma_B^2 - \sum_{i=1}^{n} (w^i)^2 \sigma_i^2}{\left(\sum_{i=1}^{n} w^i \sigma_i\right)^2 - \sum_{i=1}^{n} (w^i)^2 \sigma_i^2} \tag{4.33}$$

where σ_B, σ_i are the implied/realized volatilities of the basket and the i-th asset.

In the case that the number of assets is sufficiently large and the weights are not far from equal weights, there is a handy approximation. In equation (4.32), the order of magnitude of the first term is

$$\left(\sum_{i=1}^{n} w^i \sigma_i \right)^2 \approx \left(\sum_{i=1}^{n} \frac{1}{n} \sigma \right)^2 = \sigma^2 \tag{4.34}$$

The second term instead is

$$\sum_{i=1}^{n} (w^i)^2 \sigma_i^2 \approx \sum_{i=1}^{n} \frac{1}{n^2} \sigma^2 = \frac{1}{n} \sigma^2 \tag{4.35}$$

We clearly see that for n large enough, say 50 stocks as the number of components of the EuroStoxx 50, the second term is only 2% of the first term. Therefore it is a not too bad approximation to drop it. Hence we find

$$\sigma_B^2 \approx \rho \left(\sum_{i=1}^{n} w^i \sigma_i \right)^2 \tag{4.36}$$

Definition 4.12. The implied/realised *dirty correlation* among the n assets is defined as

$$\rho = \left(\frac{\sigma_B}{\sum\limits_{i=1}^{n} w^i \sigma_i} \right)^2 \tag{4.37}$$

where σ_B, σ_i are the implied/realized volatilities of the basket and the i-th asset.

Observation 4.13. The dirty correlation is always larger than the clean correlation. From equation (4.32) we see that $\sigma_B^2 \geq \rho_{clean} \left(\sum_{i=1}^n w^i \sigma_i\right)^2$ and the conclusion follows.

Observation 4.14. Whereas it is not possible to define an *implied* average pairwise correlation, it becomes possible to define *implied* clean/dirty correlations, provided that there are liquid options on the basket and on the underlyings.

Realized volatility of the basket

Special attention must be made when computing the realized volatility of a basket. If all the basket components are traded in the same hours, the realized volatility can be normally computed with the daily close-to-close log-returns of the basket. However, if the basket components are traded in different geographical areas, the volatility of the basket must be computed with a Picking Frequency (e.q. 3-5 days). Otherwise the time lag introduces a decorrelation among the assets and therefore reduces the volatility of the basket.

Chapter 5

Dispersion

5.1 Interest in trading correlation

Investment banks keep on selling structured products containing Worst-Of features, options on Baskets and many more sophisticated payoffs with a long exposure to correlation. On the other hand, institutional investors keep on buying Puts on indices for protection; portfolio managers keep on selling volatility on single stocks, typically through call overwriting to enhance their performance. The effect is that single stock volatility is relatively cheap and index volatility is relatively expensive. The combined behaviours of the buy side and the sell side create a spread between implied and realised correlation (which empirically is around 10 points). Investment banks find themselves with big short exposure to correlation and want to buy correlation back; on the buy side, many clients want to profit from the attractive returns offered by a short correlation exposure.
A possible way to directly trade correlation is Correlation Swaps. These instruments are simply forward contracts on the Pairwise Realised Correlation. However they are not very liquid, and there is no method to define an implied pairwise correlation. Also, there are obvious hedging problems: correlation is a very elusive quantity which is difficult to replicate with other financial products. From these reasons comes the popularity of Dispersion Trades, which provide an exposure to correlation but also benefit from the liquidity of options, they are easy to implement and easy to hedge.

5.2 General Dispersion Trade

Dispersion Trades consist in taking a position in an index option and taking the opposite position in all the index components. Let B_t be the spot price of an Equity Index (say, the EuroStoxx 50), and $S_1, ..., S_n$ be the spot prices of the n Index components. We can assume by normalization that $S_0^1 = S_0^2 = ... = S_0^n = B_0 = 1$. Let $w_1, ..., w_n$ be the weights of the components, and $\alpha_1, ..., \alpha_n$ be positive real numbers (not necessarily with unit sum). Let $O(\cdot)$ be a generic derivative product (e.g. a Call, a Put, a Straddle, a Variance Swap).

Definition 5.1. A *Dispersion Trade* "O VS O" (e.g. Call VS Call, Straddle VS Straddle etc.) on the index B with weights $\alpha_1, ..., \alpha_n$ consists in being

- Long α_i options O on the i-th component on the index
- Short 1 option O on the Index

In formulas,

$$Disp = \sum_{i=1}^{n} \alpha_i O(S_i) \; - O(B) \tag{5.1}$$

Suppose that the product O has a positive vega (which is the case for Calls, Puts, Straddles and Variance Swaps). A Dispersion trade is thus buying volatility on the single stocks and selling volatility on the basket. Since the volatility of the basket is an increasing function of the correlation, the **Dispersion Trade is short correlation.** In general, Dispersion Trades are Delta Hedged, in order to remove the directional exposure and only keep the volatility and correlation exposure.

Proposition 5.2. *If the option O is a European payoff which is a convex function of the underlying price (true for Calls, Puts, Straddles), the basket B is an arithmetic basket of the underlyings and $\alpha_i = w_i$, then the Dispersion payoff is always positive.*

Proof. By hypothesis, $O(S_i) = f(S_T^i)$ with f being a convex function. B is an arithmetic basket, i.e. $B_T = \sum_{i=1}^n w_i S_T^i$. By Jensen's inequality,

$$\sum_{i=1}^{n} w_i f(S_T^i) \geq f\left(\sum_{i=1}^{n} w_i S_T^i\right) = f(B_T)$$

$\square$

5.3 Gamma P&L

Having said that the Dispersion trades are often Delta-Hedged, it is interesting to investigate the Gamma P&L of being long the dispersion strategy.
Throughout this section, we assume that the Basket is a dynamic Basket with weights w_i, as defined in section 2.4. For the case of a static Basket with weights $\bar{w}_i$, it is sufficient to substitute

$$w_i = \frac{\bar{w}_i S_i}{B} \tag{5.2}$$

In any case, we have that the sum of the weights is one.
Let $\Gamma_1, ...\Gamma_n, \Gamma_B$ be the Gammas of the option O on the single stocks and on the Basket (assumed to be all positive), and $\sigma_1, ...\sigma_n, \sigma_B$ the volatilities. Reminding equation (2.15), we can write the Gamma P&L of the Dispersion trade:

$$P\&L^{\Gamma} = -\sum_{i=1}^{n} \frac{1}{2}\alpha_i \Gamma_i S_i^2 \left(\sigma_i^2 \Delta t - \left(\frac{\Delta S_i}{S_i} \right)^2 \right) + \frac{1}{2}\Gamma_B B^2 \left(\sigma_B^2 \Delta t - \left(\frac{\Delta B}{B} \right)^2 \right) \tag{5.3}$$

Set $\hat{\sigma}_i^2 \Delta t := \left(\frac{\Delta S_i}{S_i} \right)^2$ and $\hat{\sigma}_B^2 \Delta t := \left(\frac{\Delta B}{B} \right)^2$. To clarify, $\hat{\sigma}_i$ and $\hat{\sigma}_B$ can be seen as instantaneous realized volatilities. Using the *clean correlation* as described in definition 4.11, with ρ being the implied and $\hat{\rho}$ being the one-day realised, we can rewrite

$$\begin{aligned} \sigma_B^2 - \frac{1}{\Delta t} \left(\frac{\Delta B}{B} \right)^2 = \sigma_B^2 - \hat{\sigma}_B^2 = \sum w_i^2 \sigma_i^2 + \rho \left[\left(\sum w_i \sigma_i \right)^2 - \sum w_i^2 \sigma_i^2 \right] - \\ - \sum w_i^2 \hat{\sigma}_i^2 - \hat{\rho} \left[\left(\sum w_i \hat{\sigma}_i \right)^2 - \sum w_i^2 \hat{\sigma}_i^2 \right] \end{aligned} \tag{5.4}$$

Now we can combine the two previous equations, putting together the $\sigma_i^2 - \hat{\sigma}_i^2$ terms:

Dispersion Gamma P&L with clean correlation - dynamic Basket

$$\frac{2}{\Delta t} P\&L^\Gamma = \sum_{i=1}^{n} \left(\hat{\sigma}_i^2 - \sigma_i^2\right) \left[\alpha_i \Gamma_i S_i^2 - \Gamma_B B^2 w_i^2\right] + \\ + \Gamma_B B^2 \left\{ \rho \left[\left(\sum w_i \sigma_i\right)^2 - \sum w_i^2 \sigma_i^2 \right] - \hat{\rho} \left[\left(\sum w_i \hat{\sigma}_i\right)^2 - \sum w_i^2 \hat{\sigma}_i^2 \right] \right\}$$

Dispersion Gamma P&L with clean correlation - static Basket

$$\frac{2}{\Delta t} P\&L^\Gamma = \sum_{i=1}^{n} \left(\hat{\sigma}_i^2 - \sigma_i^2\right) S_i^2 \left[\alpha_i \Gamma_i - \Gamma_B \bar{w}_i^2\right] + \\ + \Gamma_B \left\{ \rho \left[\left(\sum \bar{w}_i S_i \sigma_i\right)^2 - \sum \bar{w}_i^2 S_i^2 \sigma_i^2 \right] - \hat{\rho} \left[\left(\sum \bar{w}_i S_i \hat{\sigma}_i\right)^2 - \sum \bar{w}_i^2 S_i^2 \hat{\sigma}_i^2 \right] \right\} \tag{5.5}$$

In the case of the number of stocks n being sufficiently large, and the weights sufficiently balanced, we can make an approximation dropping the terms in w_i^2 and thus using the *dirty* correlations, obtaining

Dispersion Gamma P&L with dirty correlation - dynamic Basket

$$\frac{2}{\Delta t} P\&L^\Gamma = \sum_{i=1}^{n} \left(\hat{\sigma}_i^2 - \sigma_i^2\right) \alpha_i \Gamma_i S_i^2 + \Gamma_B B^2 \left\{ \rho \left(\sum w_i \sigma_i\right)^2 - \hat{\rho} \left(\sum w_i \hat{\sigma}_i\right)^2 \right\} \tag{5.6}$$

where ρ is now the dirty correlation. The above P&L can be interpreted as a first **long exposure in single-stock volatility** and a second **short exposure to *covariance* (not correlation)**. In order to identify the pure exposure to correlation, we can rewrite it as

$$\frac{2}{\Delta t} P\&L^\Gamma = \sum_{i=1}^{n} \left(\hat{\sigma}_i^2 - \sigma_i^2\right) \alpha_i \Gamma_i S_i^2 + \Gamma_B B^2 \hat{\rho} \left[\left(\sum w_i \sigma_i\right)^2 - \left(\sum w_i \hat{\sigma}_i\right)^2 \right] + \\ + \Gamma_B B^2 \left(\sum w_i \sigma_i\right)^2 (\rho - \hat{\rho}) \tag{5.7}$$

or also

$$\frac{2}{\Delta t} P\&L^\Gamma = \sum_{i=1}^{n} \left(\hat{\sigma}_i^2 - \sigma_i^2\right) \alpha_i \Gamma_i S_i^2 + \Gamma_B B^2 \rho \left[\left(\sum w_i \sigma_i\right)^2 - \left(\sum w_i \hat{\sigma}_i\right)^2 \right] + \\ + \Gamma_B B^2 \left(\sum w_i \hat{\sigma}_i\right)^2 (\rho - \hat{\rho}) \tag{5.8}$$

For a rough interpretation of the previous equations, we can imagine all the σ_i to equal to a common "single-stock volatility" σ. We have

Dispersion Gamma P&L simplified to highlight the exposure to single-stock volatility and correlation

$$\frac{2}{\Delta t} P\&L^{\Gamma} = \left(\hat{\sigma}^2 - \sigma^2\right) \left[\sum_{i=1}^{n} \alpha_i \Gamma_i S_i^2 - \hat{\rho} \Gamma_B B^2\right] + \Gamma_B B^2 \sigma^2 (\rho - \hat{\rho}) \quad (5.9)$$

Dispersion Gamma P&L simplified to highlight the exposure to single-stock volatility and covariance

$$\begin{aligned} \frac{2}{\Delta t} P\&L^{\Gamma} &= \left(\hat{\sigma}^2 - \sigma^2\right) \left[\sum_{i=1}^{n} \alpha_i \Gamma_i S_i^2 - \rho \Gamma_B B^2\right] + \Gamma_B B^2 \hat{\sigma}^2 (\rho - \hat{\rho}) = \\ &= \left(\hat{\sigma}^2 - \sigma^2\right) \left[\sum_{i=1}^{n} \alpha_i \Gamma_i S_i^2\right] + \Gamma_B B^2 \left(\sigma^2 \rho - \hat{\sigma}^2 \hat{\rho}\right) \end{aligned} \quad (5.10)$$

5.4 Examples of Dispersion Strategies

5.4.1 Variance Swap VS Variance Swap Dispersion

The most classical example of a Dispersion Strategy consists in buying a Variance Swap with Vega Notional N_i on the i-th component of the index and selling a Variance Swap with Vega Notional N_B on the Basket.
Let $K_i, \hat{\sigma}_i^2$ be the Variance Swap strike and Realised Variance of the i-th component and $K_B, \hat{\sigma}_B^2$ the Variance Swap strike and Realised Variance of the Basket. The payoff of the trade is simply

$$\text{Disp} = \sum_{i=1}^{n} N_i \frac{\hat{\sigma}_i^2 - K_i^2}{2K_i} - N_B \frac{\hat{\sigma}_B^2 - K_B^2}{2K_B} \quad (5.11)$$

Set $\alpha_i = \frac{N_i}{2K_i}$ and $\alpha_B = \frac{N_B}{2K_B}$; the payoff is then

$$\text{Disp} = \sum_{i=1}^{n} \alpha_i \left(\hat{\sigma}_i^2 - K_i^2\right) - \alpha_B \left(\hat{\sigma}_B^2 - K_B^2\right) \quad (5.12)$$

Let $\hat{\rho}_{ij}$ be the realised correlation between assets i, j; we have, similarly to equation (4.29),

$$\hat{\sigma}_B^2 = \sum_{i=1}^n w_i^2 \hat{\sigma}_i^2 + 2 \sum_{i<j} w_i w_j \hat{\sigma}_i \hat{\sigma}_j \hat{\rho}_{ij} \tag{5.13}$$

We can see the strikes K_i, K_B as implied volatilities, and define the Implied Correlations ρ_{ij}. We have that

$$K_B^2 = \sum_{i=1}^n w_i^2 K_i^2 + 2 \sum_{i<j} w_i w_j K_i K_j \rho_{ij} \tag{5.14}$$

Therefore,

$$\boxed{\text{Disp} = \sum_{i=1}^n \left(\alpha_i - \alpha_B w_i^2\right) \left(\hat{\sigma}_i^2 - K_i^2\right) \; - 2\alpha_B \sum_{i<j} w_i w_j \left(\hat{\sigma}_i \hat{\sigma}_j \hat{\rho}_{ij} - K_i K_j \rho_{ij}\right)} \tag{5.15}$$

We can see the above payoff as a combination of a long volatility exposure on single stocks and a short *covariance* (not correlation) exposure $\hat{\sigma}_i \hat{\sigma}_j \hat{\rho}_{ij} - K_i K_j \rho_{ij}$.
In the case of the number of stocks being high, we can use the Dirty Correlations $\hat{\rho}, \rho$ (realised, implied) as defined in definition 4.12. From (5.12) we see that

$$\text{Disp} = \sum_{i=1}^n \alpha_i \left(\hat{\sigma}_i^2 - K_i^2\right) \; - \alpha_B \left[\hat{\rho} \left(\sum_{i=1}^n w_i \hat{\sigma}_i\right)^2 - \rho \left(\sum_{i=1}^n w_i K_i\right)^2\right] \tag{5.16}$$

The sensitivity to the volatility of the j-th component is

$$\frac{\partial \text{Disp}}{\partial \hat{\sigma}_j} = 2\alpha_j \hat{\sigma}_j - 2\alpha_B w_j \hat{\rho} \sum_{i=1}^n w_i \hat{\sigma}_i \tag{5.17}$$

Assuming that all $\hat{\sigma}_i$ are the same, we have that the Vega of the Dispersion is

$$\frac{\partial \text{Disp}}{\partial \hat{\sigma}} = 2(\sum_{j=1}^n \alpha_j - \alpha_B \hat{\rho})\hat{\sigma} \tag{5.18}$$

Weightings

There are many interesting weighting choices of the α_i.

Covariance Dispersion (in case of a small number of stocks) In order to have an as pure as possible exposure to covariance, we see from equation (5.15) that we must choose $\alpha_i = \alpha_B w_i^2$. This is suitable when the number of stocks is small; when there are many stocks, this weighting choice becomes equivalent to just selling a Variance Swap on the Basket. The payoff in this case will be

$$\text{Disp} = -2\alpha_B \sum_{i<j} w_i w_j \left(\hat{\sigma}_i \hat{\sigma}_j \hat{\rho}_{ij} - K_i K_j \rho_{ij}\right) \tag{5.19}$$

Correlation Weighted Dispersion The aim of this weighting is to be Vega Neutral, that is eliminating the exposure to volatility and only keeping the correlation exposure. From equation (5.17) and (5.18), we see that the best choice is $\alpha_i = \alpha_B w_i \hat{\rho}$ (hence the name *Correlation Weighted*). Actually, since $\hat{\rho}$ is unknown, a forecast value will be used.

Vega Weighted Dispersion The weights are $\alpha_i = w_i \alpha_B$, or alternatively $N_i = 2w_i K_i, N_B = 2K_B$ (hence the name *Vega Weighted*). This choice can be seen as a Correlation Weighted dispersion plus a long position in single-stock volatility:

$$\text{Vega Weighted} = \sum_{j=1}^{n} \alpha_B w_j (1-\hat{\rho}) \left(\hat{\sigma}_i^2 - K_i^2\right) + \text{Correlation Weighted}$$

The rationale is in the fact that correlation and volatility are positively correlated. Therefore a long position in single-stock volatility reduces the losses of the short correlation exposure when both volatility and correlation rise.

5.4.2 Synthetic Variance Swap Dispersion

Very similar to the Variance Swap Dispersion is buying α_i synthetic Variance Swaps[1] on the i-th component of the index and selling a synthetic Variance Swap on the Index.
As we saw in equation (3.21), the Gamma of the portfolio which replicates a Variance Swap is $\Gamma = \frac{2}{TS_t^2}$. So, by equation (5.9), we have

$$P\&L^{\Gamma} = \frac{\Delta t}{T} \left[\left(\hat{\sigma}^2 - \sigma^2\right) \left(\sum_{i=1}^{n} \alpha_i - \hat{\rho} \right) + \sigma^2 (\rho - \hat{\rho}) \right] \tag{5.20}$$

[1] A Synthetic Variance Swap consists in buying the replicating portfolio defined in (3.14) and Delta-Hedging it.

The P&L is in line with the case of the Variance Swap Dispersion. With the weighting choice *Correlation Weighted* $\alpha_i = w_i\hat{\rho}$, $\sum_{i=1}^n \alpha_i = \hat{\rho}$ and the P&L becomes

$$P\&L^\Gamma = \frac{\Delta t}{T}\sigma^2(\rho - \hat{\rho}) \tag{5.21}$$

As in the case of the Variance Swap, the Correlation Weighted Dispersion provides a rather pure exposure to correlation.
The *Vega Weighted* Dispersion $\alpha_i = w_i$ produces a P&L of

$$P\&L^\Gamma = \frac{\Delta t}{T}\left[\left(\hat{\sigma}^2 - \sigma^2\right)(1-\hat{\rho}) + \sigma^2(\rho - \hat{\rho})\right] \tag{5.22}$$

We again see it as a Correlation Weighted Dispersion plus a long position in single-stock volatility, which can mitigate losses when correlation and volatility surge together.

5.4.3 ATM Call VS Call Dispersion

It consists in buying α_i ATM European Calls on the i-th component of the index and selling an ATM Call on the Index, and Delta Hedging every option. Because of the liquidity of Call Options, this strategy is attractive, easy to implement and easy to hedge.
The Gamma of a Call Option in Black-Scholes model is

$$\Gamma = \frac{\Phi(d_+)}{S_t\sigma\sqrt{T-t}}$$

where σ is the volatility, $\Phi(x) = e^{-x^2/2}/\sqrt{2\pi}$ is the density function of the standard Gaussian and

$$d_+ = \frac{\left[\ln\frac{S_t}{K} + \left(r + \frac{1}{2}\sigma^2(T-t)\right)\right]}{\sigma\sqrt{T-t}}$$

At inception we can say that $\Phi(d_+)$ is roughly the same value for all the stocks and the Basket, being all the prices equal to the strike. Recall that $\sigma_B = \sigma\sqrt{\rho}$ with ρ being the dirty correlation, and prices at inception are all equal to 1. Then, by equation (5.10), we have

$$\begin{aligned}\frac{2}{\Delta t}P\&L^\Gamma &= \frac{\Phi(d_+)}{\sqrt{T}}\left\{\left(\hat{\sigma}^2 - \sigma^2\right)\left[\sum_{i=1}^n \alpha_i\frac{S_i}{\sigma} - \rho\frac{B}{\sigma_B}\right] + \frac{B}{\sigma_B}\hat{\sigma}^2(\rho - \hat{\rho})\right\} = \\ &= \frac{\Phi(d_+)}{\sigma\sqrt{T}}\left\{\left(\hat{\sigma}^2 - \sigma^2\right)\left[\sum_{i=1}^n \alpha_i - \sqrt{\rho}\right] + \hat{\sigma}^2\frac{\rho - \hat{\rho}}{\sqrt{\rho}}\right\} \quad \text{(at inception)}\end{aligned} \tag{5.23}$$

From the above formula, we see that the *Correlation Weighted* Dispersion is obtained with weight choices such that $\sum \alpha_i = \sqrt{\rho}$. However, there is a big problem in the above strategy. The Gamma of a Call Option is maximised near the strike and rapidly decreases when the spot moves away. This effect is expressed in the above formula in the $\Phi(d_+)$ term, where the Gaussian bell appears. Therefore, when spots move away from the ATM over time, the Gamma vanishes and what is captured on only noise. The strategy may be good for short maturities (e.g. 6 months) because the spots don't have much time to move away from the ATM, but for longer maturities it does not capture anymore the correlation.

Table 5.1: Pros and Cons of ATM Call VS Call

Pros	**Cons**
Uses liquid instruments Provides exposure to correlation Easy to hedge	When spots move away from the ATM, Gamma vanishes and the P&L is noise

5.4.4 Straddle VS Straddle or Put VS Put

The P&Ls of a Call Delta Hedged and of a Put Delta Hedged with same strike are equal. This can be justified by the Call-Put parity or by the fact that the Gamma of a Call is equal to the Gamma of a Put with the same strike. Therefore, provided that the Dispersion strategy is Delta Hedged, using Calls, Puts or Straddles gives the same outcome. There could be some interest in using Straddles because of their smaller delta at inception.

5.4.5 Call Strip VS Call Strip

The aim of this strategy is to deal with the vanishing Gamma issue of the ATM Call VS Call Dispersion while providing the same advantages described in table 5.1. A possible solution is to use multiple vanilla options with different strikes in order to keep the Gamma significantly high for a wider spot range.
Using too many vanillas with the widest range of strikes has the major inconvenience of delta-hedging too many options. For example, if we use 10 options and our dispersion runs on 50 components, we need to delta-hedge $10 \times (50 + 1) = 510$ options.
There are many possibilities in the choice of the weightings. One choice is,

recalling the replication of a Variance Swap, to use weights proportional to the inverse square strike $1/K^2$.
Let us choose weights such that $\sum \alpha_i = \sqrt{\rho}$ (Correlation Weighted). Assuming that the Gamma does not vanish, we see from equation (5.23) that the P&L is proportional to:

$$P\&L^{\Gamma} \propto B \cdot \frac{\hat{\sigma}^2}{\sigma} \cdot \frac{\rho - \hat{\rho}}{\sqrt{\rho}} \approx \hat{\sigma}(\sqrt{\rho} - \sqrt{\hat{\rho}}) \tag{5.24}$$

5.5 Backtest

We took an arithmetic basket of 10 stocks, whose tickers will not be disclosed for confidentiality reasons. The implied 6-month correlation of the 10 components has been around $50-60\%$ in the past 5 years. The realised 6-month correlation has almost always been below 50%, as expected (we said that usually the realised correlation is 5-10 points below the implied correlation). We chose to backtest a *Correlation Weighted* Dispersion with three 6-month vanilla options, with strikes $90\%, 100\%, 110\%$. Being the implied correlation around $50-60\%$, it seems reasonable to take weightings such that $\sum \alpha_i \approx \sqrt{\rho} \approx 80\%$. The complete strategy is therefore

$$\begin{aligned}\text{Disp} = 80\% \times \frac{1}{10}\sum_{i=1}^{10}\left(\sum_{K\in\{90\%,100\%,110\%\}} \frac{1}{K^2} Vanilla(K, S_T^i)\right) - \\ - \left(\sum_{K\in\{90\%,100\%,110\%\}} \frac{1}{K^2} Vanilla(K, B_T)\right) + \text{DeltaHedge} - \text{Premium}\end{aligned} \tag{5.25}$$

In the above equation, $Vanilla(K, S_T^i)$ is either a Call or a Put option struck at K on the i-th underlying. It does not matter whether it is a Call or a Put because the whole strategy is Delta Hedged. As can be seen in figure 5.1, the P&L of the strategy is short correlation. Moreover, the approximation in equation (5.24) manages to well describe the P&L, as shown in figure 5.2.

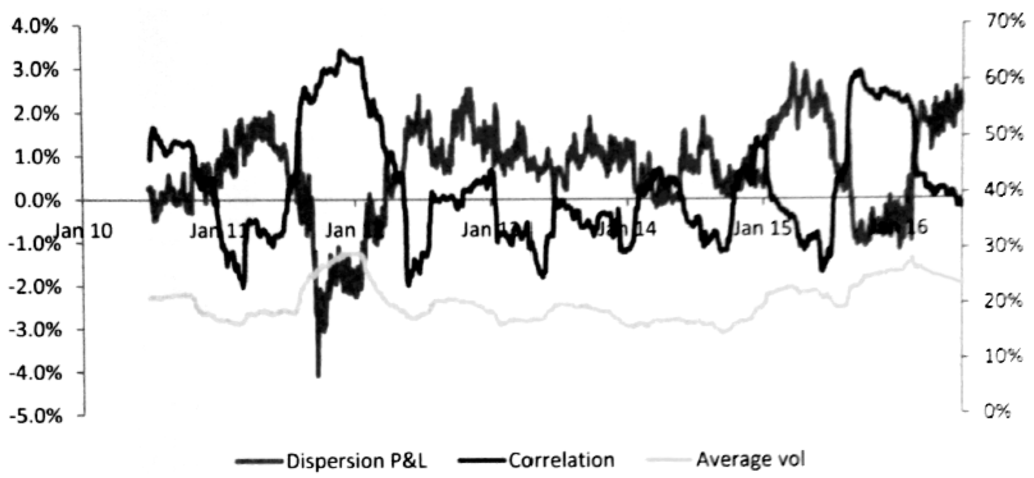


Figure 5.1: P&L of the Dispersion Strategy compared to 6-month realised correlation.

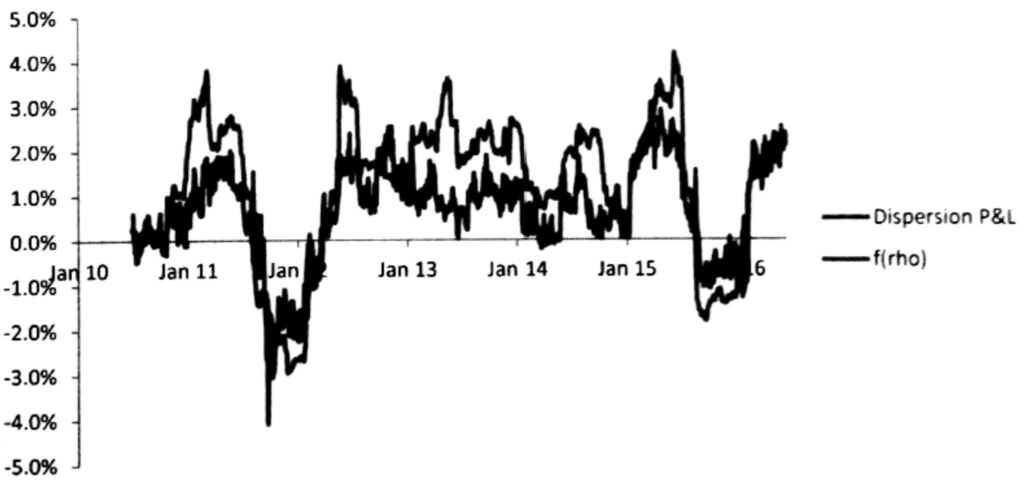


Figure 5.2: P&L of the Dispersion Strategy compared to $f(\hat{\rho}) = \hat{\sigma}(\sqrt{50\%} - \sqrt{\hat{\rho}}$. In the formula, $\hat{\rho}$ is the 6-month realised correlation and $\hat{\sigma}$ is the average of the 6-month realised volatilities of the 10 stocks.

Chapter 6

Some new ideas

The job of a Structurer also includes the creation of new, innovative products. In my opinion, this is one of the best aspects of the job: the stimulating atmosphere in the team and the outstanding skills of the colleagues spur creativity and curiosity. In this brief chapter I show two of my new ideas.

6.1 Putting together Dispersion and Variance Replication: The Gamma Covariance Swap

In this section we put together the ideas developed in the replication of Weighted Variance Swaps (subsection 3.3.3) and the Call VS Call Dispersion. The result will provide a robust replication of the Realised "Gamma Covariance", a quantity that, as far as I know, did not exist before. I chose this name because of the analogy with the Gamma Swap.

Definition 6.1. Let A_t, B_t be the prices of two assets, with volatilities σ_t^A, σ_t^B and correlation ρ_t. The *Gamma Covariance* between A_t, B_t on the period $[0, T]$ is the following quantity:

$$\Gamma Cov = \frac{252}{T} \int_0^T \frac{A_t}{A_0} \cdot \frac{B_t}{B_0} \rho_t \sigma_t^A \sigma_t^B dt \tag{6.1}$$

The *Realised Gamma Covariance* is:

$$\Gamma Cov = \frac{252}{T} \sum_{t=1}^{T} \frac{A_{t-1}}{A_0} \cdot \frac{B_{t-1}}{B_0} \log\left(\frac{A_t}{A_{t-1}}\right) \log\left(\frac{B_t}{B_{t-1}}\right) \tag{6.2}$$

The Gamma Covariance differs from the Covariance in the fact that the log-returns are weighted by the spot. The Gamma Covariance's relation

with Covariance is in a close analogy with the Gamma Swap's relation with the Variance Swap, because in Gamma Swaps the squared log-returns are spot-weighted.
In order to understand the key features of the Gamma Swap and of the Gamma Covariance, we should keep in mind the following empirical result:

Market Fact: Let A_t, B_t be the prices of two Equities, with volatilities σ_t^A, σ_t^B and correlation ρ_t. The following events usually happen either all together or none at all:

1. Both the spots A_t, B_t go down;
2. Both the volatilities σ_t^A, σ_t^B go up;
3. The correlation ρ_t goes up.

Observation 6.2. The effect of the weighting by the spot is very similar to the mitigating effect produced by the Gamma Swap. In normal market conditions (where the spots don't move too much away from their initial value), the Gamma Covariance is almost identical to the Covariance (exactly as the Gamma Swap is almost identical to the Variance Swap). In adverse market conditions, the covariance explodes: by the above Market Fact, all the three factors of the covariance=$\rho_t \sigma_t^A \sigma_t^B$ surge together. In contrast, the Gamma Covariance Swap does not explode, thanks to the mitigating effect of weighting by the spots, which are usually down in the same scenario.

6.1.1 Replication

The above observations make it clear that a Gamma Covariance Swap will have a cheaper strike than a standard Covariance Swap (as the Gamma Swap strike is cheaper than the Variance Swap strike). Another main interest of a Gamma Covariance Swap is the possibility to perfectly replicate it in the following manner:

Theorem 6.3. *The Gamma Covariance is replicable with delta strategies in the Underlyings, a zero-coupon Bond, forwards and Call VS Call (or Put VS Put) dispersions with weights* $\left(\frac{1}{4}, \frac{1}{4}\right)$ *on the single stocks and* -1 *on the* *basket. In formulas,*

$$\Gamma Cov = \frac{1}{T}\left[-2 + 2\left(\frac{\frac{A_T}{A_0}+\frac{B_T}{B_0}}{2}\right) - \frac{1}{B_0}\int_0^T \frac{A_t}{A_0}dB_t - \frac{1}{A_0}\int_0^T \frac{B_t}{B_0}dA_t - \right.$$
$$\left. - 4\int_0^{+\infty}\left[\frac{1}{4}Van_K\left(\frac{A_T}{A_0}\right) + \frac{1}{4}Van_K\left(\frac{B_T}{B_0}\right) - Van_K\left(\frac{\frac{A_T}{A_0}+\frac{B_T}{B_0}}{2}\right)\right]dK\right] \tag{6.3}$$

where $Van_K(x)$ is the payoff of the Vanilla Call/Put on x struck at K which was Out-Of-The-Money at inception, i.e.

$$Van_K(x) = \begin{cases} (x-K)_+ & \text{if } K > 1 \\ (K-x)_+ & \text{if } K < 1 \end{cases} \tag{6.4}$$

Proof. Throughout the proof, let us assume by normalization that $A_0 = B_0 = 1$.
Step 1: Ito's formula. By applying Ito's formula to the function $f(x,y) = xy$, we have

$$A_T B_T - A_0 B_0 = \int_0^T A_t dB_t + \int_0^T B_t dA_t + \int_0^T A_t B_t \rho_t \sigma_t^A \sigma_t^B dt \tag{6.5}$$

$$\Gamma Cov = \frac{1}{T}\left[A_T B_T - 1 - \int_0^T A_t dB_t - \int_0^T B_t dA_t\right] \tag{6.6}$$

We now see that, in order to replicate the Gamma Covariance, apart from implementing delta strategies, we need to replicate the European payoff $A_T B_T$. Carr's formula can't help us now, as it works only on a European payoff on a single underlying.
Step 2: Algebra trick.

$$A_T B_T = 2\left(\frac{A_T + B_T}{2}\right)^2 - \frac{1}{2}A_T^2 - \frac{1}{2}B_T^2 \tag{6.7}$$

With the above trick we decompose the payoff $A_T B_T$ as a sum of payoffs on A,B and on the basket $\frac{A_T+B_T}{2}$.
Step 3: Carr's formula. By Carr's formula (2.1) applied three times, with $S^* = 1$, we have:

$$2\left(\frac{A_T + B_T}{2}\right)^2 = 2 + 4\left(\frac{A_T + B_T}{2} - 1\right) + 4\int_0^{+\infty} Van_K\left(\frac{A_T + B_T}{2}\right)dK \tag{6.8}$$

$$-\frac{1}{2}A_T^2 = -\frac{1}{2} - (A_T - 1) - \int_0^{+\infty} Van_K(A_T)\, dK \tag{6.9}$$

$$-\frac{1}{2}B_T^2 = -\frac{1}{2} - (B_T - 1) - \int_0^{+\infty} Van_K(B_T)\, dK \tag{6.10}$$

STEP 4: COMBINE THE FIRST 3 STEPS. By summing the last three equations we obtain:

$$\begin{aligned} A_T B_T = & -1 + 2\left(\frac{A_T + B_T}{2}\right) - \\ & - 4\int_0^{+\infty}\left[\frac{1}{4}Van_K(A_T) + \frac{1}{4}Van_K(B_T) - Van_K\left(\frac{A_T+B_T}{2}\right)\right] dK \end{aligned} \tag{6.11}$$

Plugging the replication of the payoff $A_T B_T$ into equation (6.6) we find the thesis. □

Observation 6.4. With the same procedure, we might want to replicate the (not Gamma) covariance as well. However, this is not possible with the same method. We would need to apply Ito's formula to a function $G(A, B)$ such that

$$\frac{\partial^2 G}{\partial A \partial B} = \frac{1}{AB}; \quad \frac{\partial^2 G}{\partial A^2} = \frac{\partial^2 G}{\partial B^2} = 0 \tag{6.12}$$

Unfortunately, such a function does not exist.

6.1.2 Backtest

We have backtested the results in this section on the EuroStoxx 50 and the FTSE 100 (we chose this pair because they trade in the same market hours), as well as on an Equity Basket composed of 7 stocks. The backtests confirmed that

- The Gamma Covariance does not explode in the 2008 financial crisis and after the dot-com bubble (whereas the covariance does);
- The Gamma Covariance is the same as Covariance except during market stress.
- The 1-year realised Gamma Covariance is perfectly replicated by the strategy described in equation (6.3). The strategy is implemented with daily rebalanced delta strategies, and a strip of Vanilla options to reproduce the integral. We chose strikes ranging from 40% to 160%, evenly spaced by 5 points. For the ATM Dispersion, we used one half of the ATM Call VS Call and one half of the ATM Put VS Put.

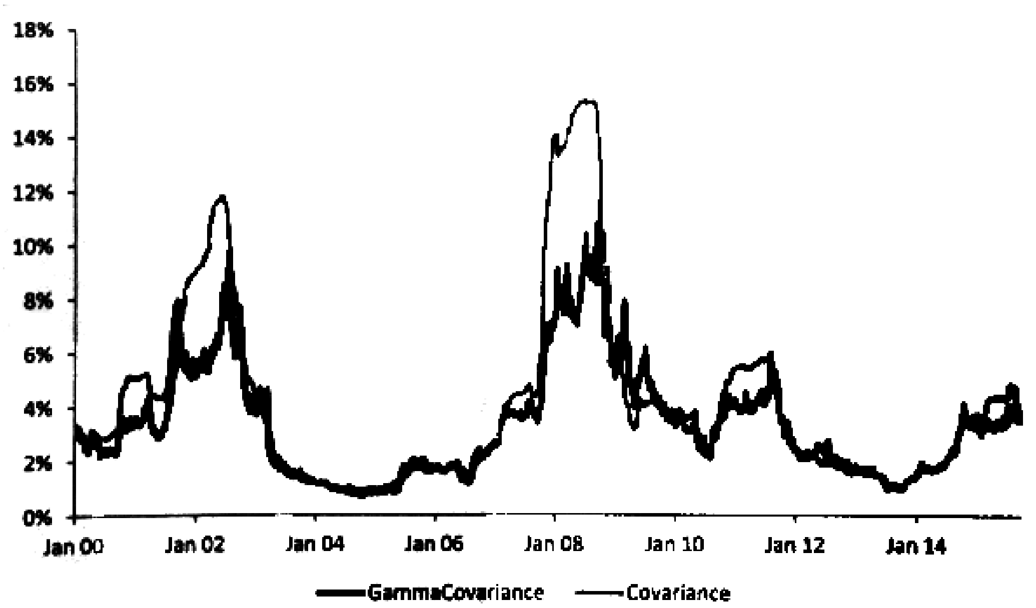


Figure 6.1: Comparison between 1-year realised Covariance and Gamma Covariance between the EuroStoxx 50 and the FTSE 100. The two quantities are almost the same in normal market conditions, but selling the Gamma Covariance brings about mitigated losses during market crashes compared to selling the Covariance.

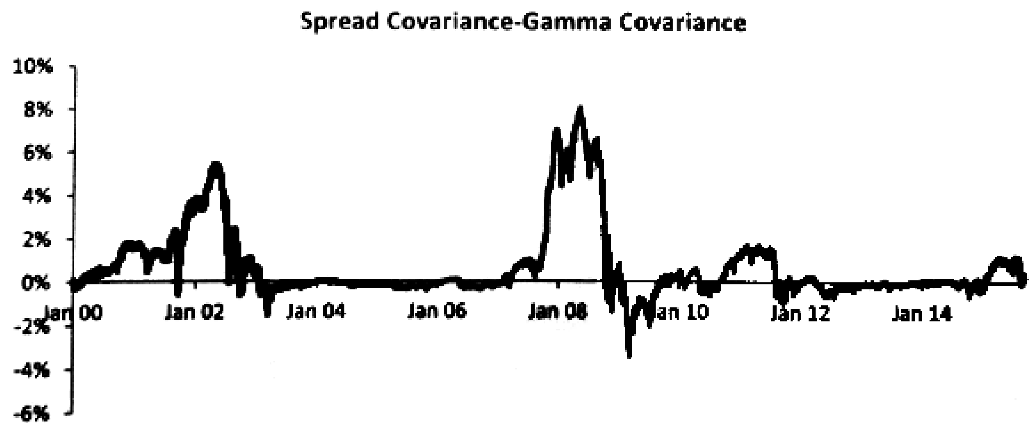


Figure 6.2: Spread between 1-year realised Covariance and Gamma Covariance between the EuroStoxx 50 and the FTSE 100.

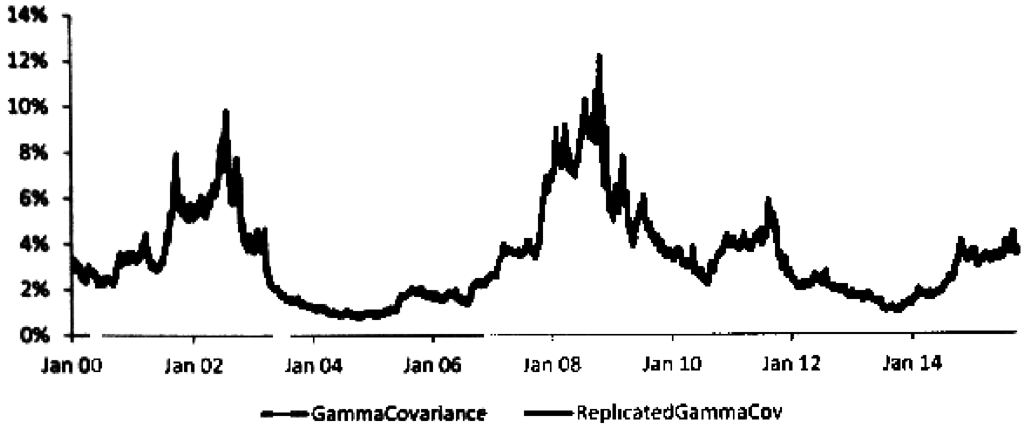


Figure 6.3: 1-year realised Gamma Covariance between the EuroStoxx 50 and the FTSE 100, and its replication according to Theorem 6.3. The replication makes use of strikes ranging from 40% to 160%, evenly spaced by 5%. The replication is almost perfect, having a replication error of 3 basis points on average.

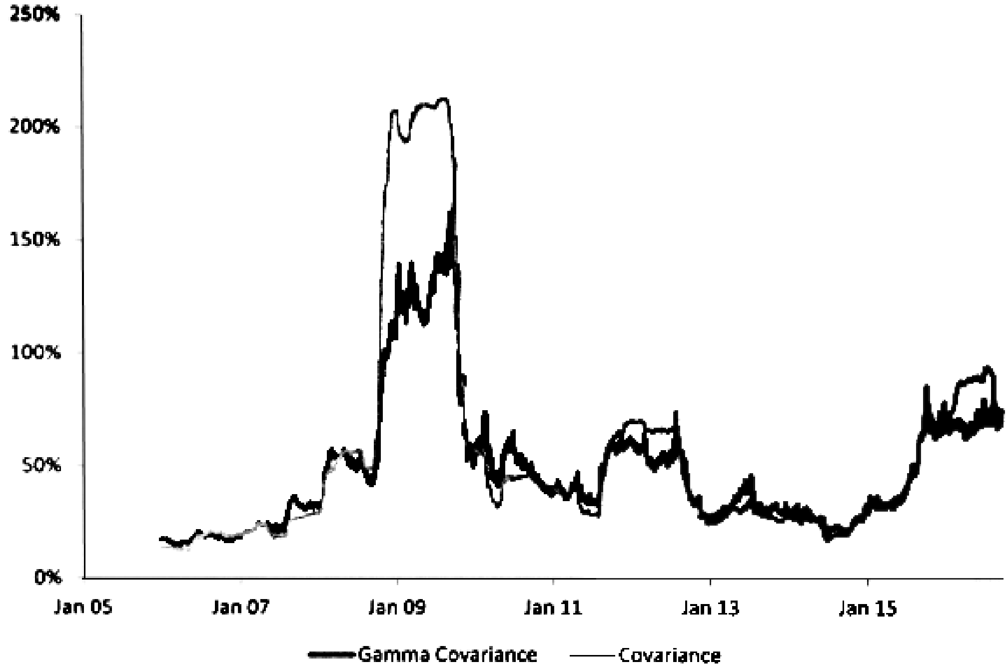


Figure 6.4: Comparison between 1-year realised Covariance and Gamma Covariance in the case of a Basket of 7 stocks. The two measures are the weighted averages of the 21 Covariances and Gamma Covariances between each pair of stocks. The weights are proportional to the inverse of volatilities at inception, so the quantities can be somehow identified with correlations (despite being larger than 1 when both correlation and volatility surged in 2008).

6.2 The FX-Gamma Covariance

The previous section is valid regardless of the asset classes of the underlying; nevertheless, it is designed and thought for two Equities. In this section I present a slightly different definition of the Gamma Covariance which is adapted to an Equity Index and a FX rate.
Let £ be a foreign currency and \$ be a domestic currency (in this example we imagine to be a US investor looking at the British Equity market, but it could be any pair of currencies). Let S_t be the price in £ of a foreign share, and X_t be the GBP/USD exchange rate (to be precise, X_t is the value in dollars of £1). Suppose that an American Depositary Receipt on S is tradable in the US market, with price $ADR_t = S_t X_t$.

Definition 6.5. The *FX-Gamma Covariance* between the foreign asset S_t and the FX rate X_t on the period $[0, T]$ is defined as

$$\mathrm{FX}\Gamma = \frac{252}{T} \sum_{t=1}^{T} \frac{S_{t-1}}{S_0} \log\left(\frac{S_t}{S_{t-1}}\right) \log\left(\frac{X_t}{X_{t-1}}\right) \tag{6.13}$$

or, in the continuous world,

$$\mathrm{FX}\Gamma = \frac{252}{T} \int_0^T \frac{S_t}{S_0} \sigma_t^S \sigma_t^X \rho_t^{SX} dt \tag{6.14}$$

where $\sigma_t^S \sigma_t^X \rho_t^{SX}$ is the covariance between S and X. Alternatively, returns could be computed with a Picking Frequency to compensate for the difference in trading hours between the two markets.

6.2.1 Replication

The FX-Gamma Covariance is very similar to the normal covariance. If we speak about an Equity and a FX rate, the covariance can be positive or negative, and tends to reach extremal values when the Equity crashes. The effect of multiplying by the price in the FX-Gamma Covariance mitigates this effect, and most importantly makes that quantity **perfectly replicable**.

Theorem 6.6. *The FX-Gamma Covariance is replicable with the following instruments: a Quanto Forward on the foreign Equity, a Delta strategy in the ADR and a Delta Strategy on the FX rate. In formulas,*

$$
\begin{aligned}
FX\Gamma = \frac{252}{T} \Bigg[& -\left(\frac{S_T}{S_0} - 1\right) + \frac{1}{S_0} \sum_{t=1}^{T} \frac{1}{X_{t-1}} \left(ADR_t - ADR_{t-1}\right) - \\
& - \frac{1}{S_0} \sum_{t=1}^{T} \frac{S_{t-1}}{X_{t-1}} \left(X_t - X_{t-1}\right) \Bigg]
\end{aligned}
\tag{6.15}
$$

Proof. Let us assume the following general diffusion model:

$$
\begin{cases}
\frac{dS_t}{S_t} = \mu_t^S dt + \sigma_t^S dW_t^S \\
\frac{dX_t}{X_t} = \mu_t^X dt + \sigma_t^X dW_t^X
\end{cases}
\tag{6.16}
$$

with $d[W_t^S, W_t^X] = \rho_t^{SX} dt$. By Ito's formula,

$$
dADR_t = d(S_t X_t) = X_t dS_t + S_t dX_t + S_t X_t \sigma_t^S \sigma_t^X \rho_t^{SX} dt \tag{6.17}
$$

Therefore, by dividing by X_t, we obtain

$$
\frac{1}{X_t} dADR_t = dS_t + \frac{S_t}{X_t} dX_t + S_t \sigma_t^S \sigma_t^X \rho_t^{SX} dt \tag{6.18}
$$

and integrating both sides provides

$$
\int_0^T \frac{1}{X_t} dADR_t = (S_T - S_0) + \int_0^T \frac{S_t}{X_t} dX_t + \frac{T}{252} \mathrm{FX}\Gamma \tag{6.19}
$$

The discrete, daily rebalanced version of the above equation multiplied by $\dfrac{252}{TS_0}$ provides the desired replication. □

6.2.2 The FX-Gamma Covariance as a form of Quanto Forward Delta Hedged and FX Hedged

We can look at the replication in theorem 6.6 in an alternative way: as a form[1] of Delta Hedge and FX Hedge of the sale of a Quanto Forward, where

[1] This form does not correspond to the true Delta Hedge and FX Hedge of a Quanto Forward, which keeps into account the Equity-FX correlation.

Deltas are computed assuming uncorrelated Equity-FX. The Quanto Forward Payoff is the amount

$$\text{Quanto Fwd} = \left(\frac{S_T}{S_0} - 1\right)(\$)$$

paid at maturity, in domestic currency. It can be rewritten as

$$\text{Quanto Fwd} = \left(\frac{1}{S_0 X_T} ADR_T - 1\right)(\$)$$

Let us compute the Deltas assuming zero Equity-FX correlation. For simplicity, we assume that S has flat forward and interest rates are zero. Under these assumptions, the price of a Quanto Forward is $E_t[S_T] = S_t$ (i.e. $E_t[\frac{ADR_T}{X_T}] = \frac{ADR_t}{X_t}$). The Delta of the Quanto Forward with respect to the ADR is

$$\frac{\partial}{\partial ADR_t}\left(\frac{1}{S_0 X_t} ADR_t\right) = \frac{1}{S_0 X_t}$$

and the Delta with respect to the FX rate is

$$\frac{\partial}{\partial X_t}\left(\frac{1}{S_0 X_t} ADR_t\right) = -\frac{1}{S_0 X_t^2} ADR_t = -\frac{S_t}{S_0 X_t}$$

The replication described in theorem 6.6 can be interpreted in the following way: what is the replication error if I sell a Quanto Forward, Delta Hedge it and FX Hedge it, but I compute the Deltas with a mistakenly zero correlation? The answer is that the residual P&L is exactly the FX-Gamma Covariance.

6.2.3 Conclusions

The FX-Gamma Covariance provides an interesting alternative to covariance. In the vast majority of cases, it is not distinguishable from covariance. Yet it is replicable with very simple instruments (whereas covariance is only replicable with more illiquid instruments like Variance Swaps) and does not explode (positively or negatively) when Equity-FX correlation approaches ± 1.

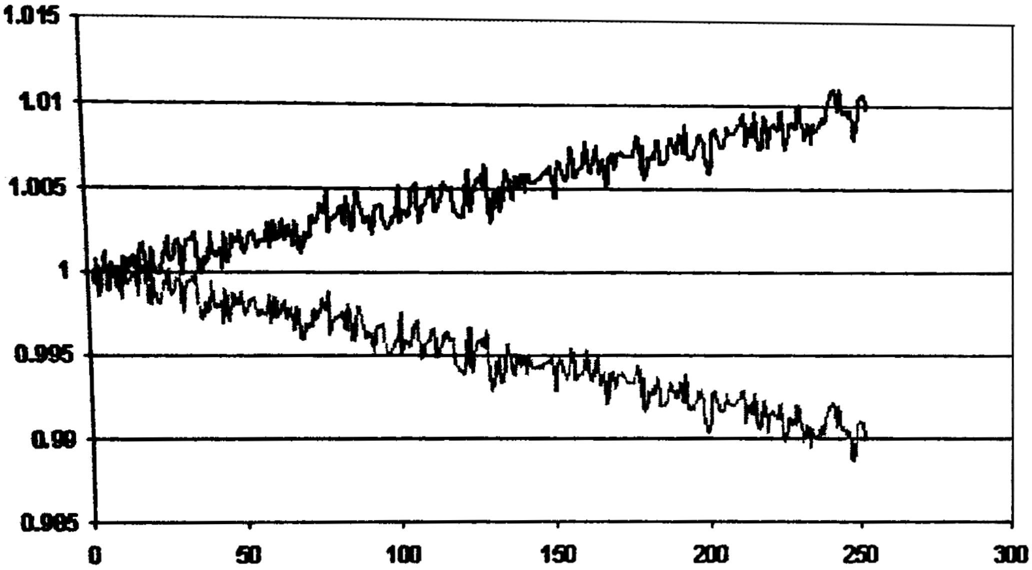
Conclusions

Volatility and Correlation are related to second order moments of random variables. In a Gaussian world, first order moments (the mean) and second order moments (variance and covariance) completely characterise the distribution of the variables. We do not live in a Gaussian world: in financial markets there are jumps, fat tails, stochastic volatility.

However, volatility and correlation remain very important key statistics to describe the behaviour of financial assets, and in particular to price structured products. Volatility and correlation have a long history and wide heritage; people are not ready yet to forget them and only use more sophisticated tools such as copulas or rank correlation. And despite the non-Gaussianity of the world, volatility and correlation remain a very simple and effective way to capture the main features of financial assets' price evolution.

Investors show a strong interest in trading these abstract quantities. While volatility has been widely traded and replicated using simpler financial products, correlation is harder to capture. Understanding which quantities are mathematically significant (covariance or correlation? Pearson or Kendall? Dirty or average pairwise?) is the first step in designing correlation structured products. When analysing Dispersion strategies, the next step is to choose the right basket, the right weights, the right option. Here Mathematics has to give more way to experience in financial markets and intuition.

Nevertheless, intuition is not always truthful. Every day we stumble upon very counter-intuitive events and Mathematics is the only way to explain what intuition misses.



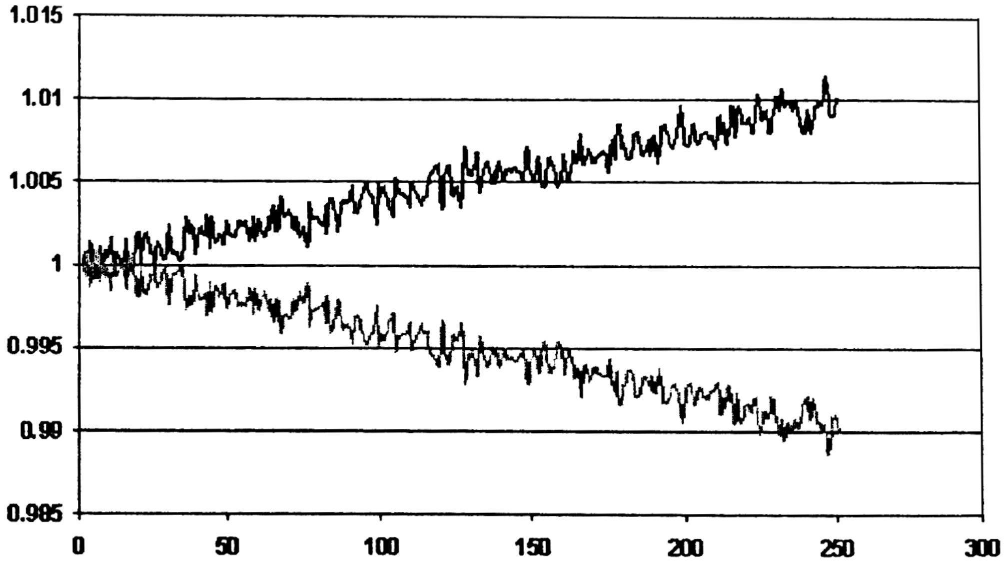


Figure 6.5: Counter-intuitive example of two trajectories having correlation +1 (top) and correlation -1 (bottom). Courtesy of Antoine Garaïalde.

Acknowledgements

Before anyone else, I would like to thank my Thesis Supervisor Maurizio Pratelli. Throughout the years, he has taught me a lot about mathematical finance and provided invaluable advice. He has been very kind and helpful, especially because of the unusual situation where we collaborated by emails, and I appreciate the time he spent writing me and helping me with the forms and documents.

Heartfelt thanks also to my colleagues in J.P. Morgan. First of all, my Guru Antoine Garaïalde: every day he teaches me something new. He involved me in thrilling projects about various Volatility and Correlation trades, gave me advice and stimulated my curiosity. All of that with a good amount of good mood and laughter.
I want to thank Arnaud Jobert, for the trust he gave me since the beginning, for the way he makes me feel welcome and for the amazing opportunity to join his team.
Thanks to Guillaume Flamarion, who shares with me his impressive experience and knowledge of the job and helps me gain visibility within the firm.
Thanks also to Andrea Santacroce, who is the main reason why I am in this amazing Structuring team in J.P. Morgan. From the moment we met in Paris, he opened my eyes on my career ambitions and gave me invaluable advice. And since the start of my position in J.P. Morgan, he dedicated huge amounts of his time to help me, teach me, guide me.
Thanks to Johannes Ingerfurth for teaching and helping me as a friend does.
Many thanks to all the rest of the Structuring and Trading team, especially to Julien Nizard, for their help and guidance and for making me happy every day to go to work.

Thanks to my mum, for her fond love and support and for making me always feel special. Despite the distance, she is always ready to pick up the phone and make us feel close.
Thanks to my dad for his high consideration towards me, for his love and for visiting me in London.
Thanks to my brother, for not visiting me in London :) and for always making me laugh and feel like a kid.

Heartfelt thanks to Cristofer. Literally, without him I would have never graduated. He did so much for me. He came to London many times and kept me smiling in the hardest situations. He supported me every day at the phone. He would do anything to help me, and he did help me so much indeed.

Thanks to all my friends, for always keeping in touch despite the distance, for their love and support.
Thanks to my friends from Rome, for their long-lasting friendship: Patonza, Dark Lady, Bionda, Lumière, Piccola Panda, Foca, Cagna, Piva.
Thanks to my friends in Pisa, who made my four years in Tuscany really amazing: Gatta, Vicky, Pamela, Alessia, Erika, Marsha&Labrador, Ila, Orsetto&Enrico, Terranova, Lemurino&Cigno, Megumi&Psycho, Davide, Chiaretta, Ghioncé, le Gallinelle (Serena, Angelica, Francesca, Chiara, Giulia), Marti, Giulia la Rossa, Luca, Marco, Micky, Anto, Elena&Lori, il Principe, Dani and everybody else (there would be no space here to list you all!).
Thanks to my flatmate Mathieu for always keeping me in good mood and thanks to Silvia for great nights in London and her new promising friendship!!

Bibliography

[1] D. Revuz, M. Yor, *Continuous Martingales and Brownian Motion*, 1991

[2] T. Björk, *Arbitrage Theory in Continuous Time*, 2004

[3] J. Hull, *Options, Futures and Other Derivatives*, 2009 (Seventh Edition)

[4] R. Lee, *Weighted Variance Swap*, 2009

[5] K. Demeterfi, E. Derman, M. Kamal, J. Zou, *More Than You Ever Wanted To Know About Volatility Swaps*, Goldman Sachs Quantitative Research, 1999